

Foundations of a Nonideal Theory of Justice

Marcus Arvan
University of Tampa

Political theorists are increasingly interested in the relationship between "ideal theory" and "nonideal theory." A number of influential political theories – including John Rawls' theory of domestic justice – focus primarily on describing ideal, or fully just, social conditions.¹ Rawls constructed a novel version of the idea of a social contract, the "original position," in order to derive principles of justice. However, Rawls only applied the original position to the case of a "well-ordered society" – one in which, "(1) everyone accepts and knows that the others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are generally known to satisfy these principles."² Rawls recognized this to be a great idealization. He even recognized that because of it, his principles of justice and their priority relations may not apply at all under less ideal conditions.³ Unsurprisingly, the question of how to extend Rawls' theory to nonideal conditions (or nonideal theory) has received much attention. Some argue that Rawls' theory is overly ideal and cannot be extended to nonideal conditions.⁴ Some even argue

¹ See John Rawls, *A Theory of Justice: Revised Edition*. (Cambridge, MA: The Belknap Press of Harvard University Press, 1999a), 4-5, 8-9, 215-6. Other theories that focus primarily on ideal theory include, but are by no means limited to Charles Beitz, *Political Theory and International Relations*. (Princeton: Princeton University Press, 1979, 1999); Simon Caney, *Justice Beyond Borders: A Global Political Theory*. (Oxford University Press, 2005); Andrew Kuper, *Democracy Beyond Borders: Justice and Representation in Global Institutions* (Oxford: Oxford University Press, 2004); Darrel Moellendorf, *Cosmopolitan Justice*. (Oxford: Oxford University Press, 2002); Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974); and Thomas Pogge, "An Egalitarian Law of Peoples." *Philosophy and Public Affairs*, 23 (1994), 195-224.

² Rawls (1999a): 4-5, 8-9, 215-6.

³ Rawls (1999a): 215-6 wrote, "The principles and their lexical order were not acknowledged with these situations in mind and so it is possible that they no longer hold."

⁴ See Katherine Eddy, "Against Ideal Rights." *Social Theory and Practice*. Vol. 34, No. 3 (2008); Anne Phillips. "Egalitarians and the Market: Dangerous Ideals," *Social Theory and Practice*, Vol. 34,

that ideal theory inherently overidealizes, and that political theory should focus exclusively on nonideal theory.⁵ Rawlsian orthodoxy, however, has long held that Rawls' theory *can* be extended to nonideal theory. The orthodox view (as expressed by Korsgaard, Phillips, Stemplowska, Taylor and others⁶) is Rawls' principles and their priority relations should be extended to nonideal conditions in "spirit." Recently, this orthodox view has come under attack. John Simmons and a few others argue that because Rawls entirely idealizes away from the kinds of transition costs people can expect to face under nonideal conditions, some account of *transitional* fairness is necessary to properly extend Rawls' theory to nonideal conditions.⁷ Simmons never gives an analysis of transitional fairness, however, and no such theories exist to date. Moreover, Simmons is skeptical that such a theory is possible. He writes, "we should not expect anything less sloppy" than "somewhat speculative" extensions of Rawls' ideal theory to nonideal theory.⁸

This paper uses Rawls' original position to go far beyond Simmons' critique of Rawlsian orthodoxy. I show, against Rawlsian orthodoxy and beyond Simmons, that free and equal citizens who accept Rawls' principles of ideal theory and their priority relations should prioritize a new class of *nonideal theoretic primary goods* over the

No. 3 (2008); and Amartya Sen, *The Idea of Justice*. (Cambridge, MA: The Belknap Press of Harvard University Press, 2009).

⁵ See Colin Farrelly, "Justice in ideal theory: A refutation." *Political Studies* (2007); and Sen (2009).

⁶ See Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996): 147-151; Michael Phillips, "Reflections on the Transition from Ideal to Non-Ideal Theory," *Nous*, Vol. 19, No. 4 (1985); Zofia Stemplowska, "What's Ideal About Ideal Theory?," *Social Theory and Practice*, Vol. 34, No. 3 (2008); and Robert S. Taylor. "Rawlsian Affirmative Action," *Ethics*, 119 (2009): 476-506, esp. 485-492.

⁷ See A. John Simmons, "Ideal and Nonideal Theory." *Philosophy and Public Affairs*, 38, no. 1 (2010): 5-36.

⁸ *Ibid*: 19-20.

satisfaction of those principles and priority relations. These nonideal theoretic primary goods are all-purpose means for accomplishing all three of the following three goals:

- (A) Causing people who do not accept Rawls' principles and priority relations as an aim to come to have those principles and priorities as an aim;
- (B) Enabling people who do have Rawls' principles and priority relations as an aim to rationally weigh those principles and priority relations against any and every type of personal cost they might face in nonideal conditions; and
- (C) Enabling people who have Rawls' principles and priorities as an aim to effectively pursue their favored weighting of Rawls' principles and priority relations against such costs.

I then show that at least three specific "goods"⁹ are all-purpose means of this sort (and therefore qualify as nonideal-theoretic primary goods):

- i. Opportunities to participate equitably and effectively in a (broadly¹⁰) Rawlsian grass-roots reform movement;
- ii. General education, by various methods (political speech, pamphleteering, etc.), in Rawlsian ideals and information relevant to both (a) rationally weighing Rawlsian ideals against personal costs, and (b) effectively pursuing one's favored weightings under the particular nonideal conditions in which one lives; and,

⁹ Some of the nonideal-theoretic goods I describe – in particular, the good of having social progress conform to one's "reformist preferences" – are not goods in the "normal" sense: they are not tangible things like money or property. But this is not unique to my account. Rawls' own list of (ideal-theoretic) primary goods included all sorts of non-tangible things: rights and liberties, opportunities, and social props to self-respect. See Rawls (1999a).

¹⁰ I will say much more about this parenthetical issue in §5.

- iii. Having the course of social reform conform to the collective decisions of the (broadly) Rawlsian reform group of which one is a member.

Finally, I show that free and equal citizens would agree to a general principle for distributing nonideal-theoretic primary goods, and by extension, three corollary principles, each of which takes lexical priority over those that follow. The general principle is:

General Principle of Rawlsian Nonideal Theory: For any class of unjust conditions U , the more a given population P is unjustly disadvantaged under U , the more goods for (A) causing others to come to have Rawlsian ideals, (B) enabling those who have Rawlsian ideals to weigh those ideals against personal costs, and (C) enabling those who weigh Rawlsian ideals against personal costs to successfully pursue their favored weightings, the members of P ought to enjoy.

The three corollary principles then are:

First Corollary (The Grass-Roots Principle): The more a given population P is unjustly disadvantaged under nonideal conditions U , the more opportunities the members of P ought to have to participate *equitably and effectively in a (broadly) Rawlsian grass-roots social movement*.

Second Corollary (The Education Principle): The more a given population P is unjustly disadvantaged under nonideal conditions U , the more the members of P ought to have the common aim of *educating* people in: (a) Rawlsian ideals, (b) skills and information relevant to rationally weighing Rawlsian ideals against

personal costs, and (c) skills and information relevant to effectively pursuing such weightings.

Third Corollary (The Grass-Roots-Decision Principle): The more a given population P is unjustly disadvantaged under U , the more the overall course of social reform ought to conform to the *collective decisions* of grass-roots movements that satisfy the first Two Corollaries.

I aim to show all of this by beginning a long overdue project, one that Rawls once alluded to but which has never been attempted: systematically extending his original position to nonideal theory.¹¹

§1 of this paper gives a *prima facie* argument for extending the original position to nonideal theory. §2 then shows that if Rawls' arguments in *A Theory of Justice* and *Political Liberalism* are correct, people only need to have Rawls' principles and priority relations as *an* aim under nonideal conditions, not their only aim or even their predominant aim. Indeed, I show that if Rawls' arguments are correct, each person has an equal right to *weigh* Rawls' principles and priorities against whatever personal costs they might incur (as individuals) in the course of social progress. §3 concludes, as such, that Simmons' critique of Rawlsian orthodoxy does not go nearly far enough. Because each person has an equal right to weigh Rawls' principles and priorities against personal costs under nonideal conditions, and a nonideal-theoretic stage of the original position models this very idea, we cannot know what principles of Rawlsian nonideal theory there are until we completely analyze the *nonideal-theoretic original position*. §4 then turns to

¹¹ Rawls (1999a) alludes to this on p. 216, where he writes that it is only after principles of ideal theory are selected that, "the parties ask which principles to adopt under less happy conditions."

this task, describing its “veil of ignorance” and its parties’ overall deliberative situation. §5 argues that, given their situation, the parties should pursue the three nonideal-theoretic primary goods discussed above (as well as certain priority relations between those goods). Finally, §6 argues for the General Principle of Rawlsian Nonideal Theory and, by extension, the Three Corollaries. I provide two related arguments for the General Principle. §6.1. argues that Rawls’ famous “maximin argument” not only leads to the General Principle, but indeed, is *far more immune to criticism* within nonideal theory than (Rawls’ case for the rule is) within ideal theory.¹² §6.2. then argues that the General Principle and its Three Corollaries fair exceedingly well in “reflective equilibrium” – the process of testing them against our considered moral judgments.

Three caveats are in order before we begin. First, this paper only discusses Rawls’ theory of domestic justice (and, by extension, “Rawlsian Cosmopolitan” theories that recast Rawls’ domestic theory as a theory of global justice¹³). I leave discussion of Rawls’ international theory for another day.¹⁴ Second, in constructing a nonideal theoretic stage of the original position, I assume that Rawls’ arguments in *A Theory of Justice* and *Political Liberalism* are successful. Although I believe my case for the General Principle of Rawlsian Nonideal Theory succeeds without assuming the success of Rawls’ ideal theoretic arguments, my primary aim here is to extend Rawls’ theory to nonideal theory as-is. Finally, although Rawls’ arguments in *A Theory of Justice* and

¹² It is crucial to note that over time, consensus has arisen that, “maximin reasoning plays no role in the argument for [Rawls’] difference principle.” (See the Stanford Encyclopedia of Philosophy Entry, “Rawls”, §4.7.). Indeed, Rawls himself only cites the rule as a “useful heuristic” (Rawls 1999a: 132). My argument for maximin within nonideal theory is in no way predicated upon Rawls’ use (or non-use) of the rule within ideal theory).

¹³ See e.g., Beitz (1999), Caney (2005), Kuper (2004), Moellendorf (2002), and Pogge (1999).

¹⁴ See John Rawls, *The Law of Peoples*. (Cambridge, MA: Harvard University Press, 1999b).

Political Liberalism are strictly incompatible as a whole (Rawls took *Political Liberalism* to correct some errors in *A Theory of Justice*)¹⁵, §2 demonstrates how the portions of Rawls' arguments from *A Theory of Justice* and *Political Liberalism* that are relevant to our inquiry converge.

§1. A *Prima Facie* Case for Constructing a Nonideal Theoretic Original Position

Rawls claims that “justice is fairness,” and that the original position is a fair procedure for deriving principles of justice. Yet Rawls only applies the original position to “well-ordered” conditions of “strict compliance” – conditions in which “(1) everyone accepts and knows that the others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are generally known to satisfy these principles.”¹⁶ Due to this restriction, Rawls recognizes that his arguments at most establish what a fully just society would be like, not what fair and just responses to injustice are:

It will be recalled that strict compliance is one of the stipulations to the original position; the principles of justice are chosen on the supposition that they will be generally complied with. Because the parties are choosing a conception of justice suitable for favorable conditions...the principles [of justice] define then a perfectly just scheme; they belong to ideal theory and set up an aim to guide the course of social reform. But even granting the soundness of these principles for this purpose, we must still ask how well they apply to institutions under less than favorable conditions, and whether they provide any guidance for instances of

¹⁵ See Rawls (1993): xviii-xxxii.

¹⁶ Rawls (1999a): 4.

injustice. The principles and their lexical order were not acknowledged with these situations in mind and so *it is possible that they no longer hold*.¹⁷

Rawls also wrote of a method for dealing with this issue:

The intuitive idea is to split the theory of justice into two parts. The first of ideal part assumes strict-compliance and works out the principles that characterize a well-ordered society under favorable conditions. It develops the conception of a perfectly just basic structure and the corresponding duties and obligations of persons under the fixed constraints of human life...Nonideal theory, the second part, is worked out after an ideal conception of justice has been chosen; only then do the *parties* ask which principles to adopt under less happy conditions.¹⁸

The implicit idea here – that the original position must be used to determine principles of nonideal theory – follows from two simple ideas. If we assume (a) that justice is fairness, and (b) that the original position models a fair deliberative process, then it is only by (c) applying the original position to nonideal conditions that we determine what is fair (and just) under nonideal conditions.

Now, it is hard to say why no one has pursued this project. Rawls thought that nonideal theory might be too complex or muddled to submit to systematic analysis.¹⁹

James Woodward, on the other hand, worried that a nonideal theoretic social contract might be self-contradictory.²⁰ Yet neither actually gave an argument for these views.

¹⁷ Ibid: 215-6; emphasis added.

¹⁸ Ibid: 216; emphasis added.

¹⁹ Ibid: 267. Also see John Rawls, *A Theory of Justice*. (Cambridge, MA: The Belknap Press of Harvard University Press, 1971): 303.

²⁰ James Woodward, “Commentary: Liberalism and Migration” in Brian Barry and Robert E. Goodin, eds, *Free Movement: Ethical issues in the transnational migration of people and of money* (University Park, Pennsylvania: The Pennsylvania State University Press, 1992): 78-9.

The question of whether the original position can be systematically extended to nonideal theory is at this point an entirely open question—one that I will now answer.

§2. A Nonideal Theoretic Original Position – Part I: The Parties’ Motivations

In order to extend the original position to nonideal theory, we must address the fact that Rawls developed his ideal theory in two subtly different ways in *A Theory of Justice* and *Political Liberalism*. Let us begin with Rawls’ arguments from *A Theory of Justice* and then turn to *Political Liberalism*.

§2.1. *A Theory of Justice*

In *A Theory of Justice*, Rawls argues that we should understand principles of justice as deriving from a fair hypothetical agreement among free and equal citizens. Rawls contends that in order to ensure fairness, all should have to deliberate in complete ignorance of their personal characteristics: their race, gender, religion, conceptions of the good, and so on.²¹ This is the “veil of ignorance.” Although Rawls ultimately provides a detailed “thin theory of the good”,²² as far as the original position is concerned, Rawls assumes the traditional tenets of game theory. The parties are to be conceived as purely instrumentally rational, aiming to successfully satisfy their preferences whatever their preferences are (“The concept of rationality must be interpreted as far as possible in the narrow sense, standard in economic theory, of taking the most effective means to given ends”).²³

²¹ Rawls (1999a): §4 and Chapter III.

²² Ibid: §60.

²³ Ibid: 12.

These two assumptions – the veil of ignorance and Rawls’ conception of the parties’ rationality – pose a deliberative problem. The problem is this: how can the parties deliberate instrumentally if, thanks to the veil of ignorance, they have no knowledge of who they are in society, and hence have no knowledge of whose preferences are in fact theirs? Rawls rejects one response to this problem: utilitarianism. Although Rawls rejects utilitarianism without referring directly to the original position, his argument can be understood and clarified from that perspective. Consider, once again, the parties’ situation. Because they are behind the veil of ignorance, they do not know whose preferences are in fact theirs. If they suppose that every person they could turn out to be (once the veil is raised) wants to be happy above all else, then the parties would maximize their chances of being happy by agreeing to utilitarianism. Rawls rejects utilitarianism, however, because it does “not...take seriously the plurality and distinctness of individuals.”²⁴ Indeed, while utilitarianism might maximize their *chances* of being happy, it also gives rise to the distinct possibility that they will end up being someone whose happiness must be *sacrificed* for the greater happiness of the many. This would be a disastrous result. The parties should therefore aim to avoid the possibility of such sacrifice, if possible. Rawls’ answer is that they *can* avoid this possibility by pursuing “social primary goods”: basic rights and liberties, opportunities, income and wealth, and social props to self-respect. By pursuing these are “all purpose” goods, the parties assure themselves that they will be able to successfully pursue their ends no matter who they turn out to be. As Rawls explains,

²⁴ Ibid: 26.

Regardless of what an individual's rational plans are in detail, it is assumed that there are various things which he would prefer more of rather than less. With more of these goods men can generally be assured of greater success in carrying out their ends, whatever these ends may be. The primary social goods, to give them in broad categories, are rights, liberties, opportunities, and income and wealth. (A very important primary good is a sense of one's own worth...)²⁵

Thus, even though the parties are deprived of information about their particular ends, they have enough knowledge to rank alternatives. They know that in general they must try to protect their liberties, widen their opportunities, and enlarge their means for promoting their aims whatever these are.²⁶

Rawls' argument is that the parties should agree to distribute these goods according to his two principles of ideal theory (and those principles' priority relations). The nature of primary goods is of crucial importance, however, as they pose a problem for extending his principles (and their priority relations) directly to nonideal theory.

Consider the kinds of personal costs that people often have to face under nonideal conditions for the sake of obtaining Rawlsian primary goods. Here is a timely example: in June of 2009, the Islamic Republic of Iran held an election thought by many people to be fraudulent. In the aftermath of the election, the Iranian government violently quelled protests and restricted free speech. Now, according to Rawls' ideal theory, this situation was profoundly unjust. People did not enjoy the basic rights and liberties that they should have according to his first principle of justice. If they had those basic rights and

²⁵ Ibid: 79

²⁶ Ibid: 123.

liberties, they would generally be in a better position to pursue their ends, whatever their ends are. Yet *coming* to obtain such goods is rarely a costless endeavor. Social progress, even when it moves in a generally positive direction, imposes personal costs on people. Someone who is killed during social progress, obviously, never enjoys the primary goods they are due. Other costs are less dire: threats, harassment, social instability, strife, etc. Social progress typically imposes significant costs on people. Now, because Rawls entirely idealized away from nonideal conditions, he never gave any account of how these types of costs should be weighed against his principles and their priority relations. Consequently, the parties to a nonideal theoretic stage should regard Rawls' principles and priority relations only as *an* aim of theirs, one that they should be *entirely* free to weigh against personal costs they might face under nonideal conditions. Interestingly, this appears to be precisely the result that Rawls accepted himself – for he wrote, “the principles set up *an* aim to guide the course of social reform.”²⁷

This, however, requires some qualification. Because the parties previously agreed to Rawls' principles and their priority relations for particular *reasons* within ideal theory, the parties must consider those *same* reasons within nonideal theory. The parties should weigh Rawls' principles and priority relations only against personal costs of comparable importance. To illustrate, no one who accepts Rawls' first principle and its priority over his second principle for the reasons Rawls provides within ideal theory (see Rawls' remarks on self-respect²⁸) will think that relatively minor personal costs (e.g. mere inconvenience) outweigh that principle and its priority under nonideal conditions. Still, the parties should acknowledge that some costs they might face under nonideal

²⁷ Ibid: 215.

²⁸ Ibid: 475-479.

conditions *are* serious enough to outweigh the full satisfaction of Rawls' first principle. One such cost is their life or the lives of their loved ones. Another such cost concerns self-respect. People who desire equal basic liberties may wish to avoid processes of social change that they find degrading, even if those processes result in conformity with Rawls' first principle.

Of course, different people are likely to disagree about which sorts of personal costs are significant enough to weigh against Rawls' principles and their priority relations. This being the case, I submit that the best way to proceed is to refrain from imposing any substantial restrictions on the sorts of weighings the parties to the nonideal theoretic original position can engage in. Why, one might ask, should we err on the side of permissiveness? I propose that we adopt the permissive view for now and evaluate it in "reflective equilibrium" (as Rawls claims we should of all political principles).²⁹ If, by adopting the permissive approach, we are led to principles of nonideal theory that are incompatible with our considered convictions, then perhaps we should revise the procedure, building new restrictions in. All the same, the permissive approach will be shown to justify a compelling general principle of nonideal theory.

It may appear unclear at this point how the parties could proceed in their deliberations if they lack any substantial views about which sorts of personal costs they can legitimately weigh against Rawls' principles and their priority relations. We will discuss this problem later (in §5), but I will say for now that I address it similarly to the way that Rawls tackles an analogous problem within ideal theory: by recourse to primary goods (in our case, *nonideal theoretic* primary goods). For now, it is important to

²⁹ Ibid: 18-9, 42-5, 507-8.

reiterate how weak the implications of Rawls' ideal-theoretic arguments from *A Theory of Justice* are for non-ideal theory. Because Rawls entirely idealized away from nonideal conditions, the parties to a nonideal theoretic stage of the original position must all have an equal right to *weigh* Rawls' principles of ideal theory and their priority relations against *whichever* personal costs they might face under nonideal conditions. We will now see that Rawls' arguments in *Political Liberalism* lead to the same conclusion.

§2.2. Political Liberalism

In *Political Liberalism*, Rawls argues that the original position represents our “two moral powers”³⁰ as free and equal citizens: our powers to be rational³¹ and reasonable.³² Rawls claims that these two powers entail that we have three higher-order interests.³³ The first two higher-order interests are derived from the first moral power: rationality. According to Rawls, rational individuals aim to effectively pursue their actual or present goals in life. Whatever one's actual plans are, it is rational to pursue them effectively, not ineffectively. Call this the *Interest in One's Actual Life Plans*. Second, Rawls argues that rational individuals have an interest in reflecting upon, revising, and pursuing new life plans. Rational people routinely ask themselves whether their present goals and life plans are worthwhile, and if they are not worthwhile, they pursue new ones. Call this the *Interest in Rethinking and Revising One's Plans*. Finally, Rawls claims that *reasonable* people, as free and equal citizens, should understand and aim to uphold fair principles of social cooperation. The parties to Rawls' ideal theoretic stage of the original position do

³⁰ Rawls (1993): 19, 83, 103, 108.

³¹ Ibid: 50-1.

³² Ibid: 43, 48, 54, 58, 81.

³³ Ibid: 74-5, 106.

not pursue this higher-order interest; rather, it is embodied by the original position's veil of ignorance.³⁴ Call this the *Interest in Being Moved By Justice*.

Rawls never argues for any priority relations among these higher-order interests. Indeed, what would such an argument look like? Consider a person under nonideal conditions. This person may want society to become more just (in conformity with her *Interest in Being Moved By Justice*). Second, she may realize that Rawls' principles of ideal theory place her in a better overall position to rethink, revise, and pursue new life plans (in conformity with her *Interest in Rethinking and Revising One's Plans*). Finally, even though both of these interests are important, she might not want her actual life plans upset by social progress, even if it is in a positive direction (in conformity with her *Interest in One's Actual Life Plans*). Her first two higher-order interests thus conflict with the third under nonideal conditions. Must she prioritize the first two over the third? Given that all three interests are *equally* founded in her two moral powers, it is hard to see why she must. Having the two moral powers as a free and equal citizen, she should be free to order her higher-order interests as *she* wishes, provided that her doing so is consistent with the equal right of all other citizens to order their higher-order interests as they wish. But this is identical to the proposition that free and equal citizens in a nonideal theoretic original position all have an *equal right to weigh* Rawls' principles and priority relations against whichever personal costs they might face under nonideal conditions. This is the conclusion we reached in §2.1. (in the case of *A Theory of Justice*).

³⁴ See Rawls (1993): 103f.

§3. Beyond Simmons' Critique of Rawlsian Orthodoxy: Why Rawlsian Nonideal Theory Requires a Complete Analysis of the Nonideal-Theoretic Original Position

Rawlsian orthodoxy has long held that Rawls' principles and priority relations can be extended to nonideal theory relatively straightforwardly. Michael Phillips was perhaps the first to defend this view, arguing that the proper aim of (Rawlsian) nonideal theory is to promote the *values* of (Rawlsian) ideal theory under nonideal conditions.³⁵ Christine Korsgaard has also influentially defended this idea, arguing that Rawlsian nonideal theory must respect (1) Rawls' "general" conception of justice, (2) the priority relations among Rawls' principles, and (3) the "spirit" of Rawls' ideal theory.³⁶ Stemplowska similarly claims that Rawls' theory operates simultaneously at an ideal and nonideal level.³⁷ And Robert S. Taylor uses Korsgaard's analysis to argue that Rawls' theory prohibits affirmative action quotas within nonideal theory.³⁸

John Simmons has, as I have mentioned, provided an influential critique of this view. Simmons points out, as I do here, that because Rawls assumes strict-compliance within ideal theory, Rawls has provided no analysis at all of transitional fairness, and so no analysis of how well (if at *all*, to paraphrase Rawls again) Rawls' principles and priority relations hold under nonideal conditions.³⁹ Yet Simmons says little more than this. He never purports to give an analysis of what a Rawlsian analysis of transitional fairness might look like, and again, he is skeptical that such an analysis can be

³⁵ See Phillips (1985): 565.

³⁶ Korsgaard (1996): 147-151.

³⁷ See Stemplowska (2008): §2.

³⁸ Taylor (2009), esp. 485-492.

³⁹ See Simmons (2010): 5-36.

provided.⁴⁰ It is now clear, given our analysis in §2, that Simmons' analysis is incomplete. A systematic analysis of fairness is in fact possible.

We have seen that every individual under nonideal conditions has an *equal right* to weigh Rawls' principles and priority relations against personal costs (and pursue their most favored weightings). But what exactly does it mean to say that each person has such a right? If justice is fairness, then each person has an equal right to weigh Rawls' principles (and such) *subject to conditions of fairness*. But this is exactly what the nonideal-theoretic original position models. Until we provide a complete analysis of the nonideal-theoretic original position, we have *no* analysis of what is fair and just under nonideal conditions. If justice is fairness, and the original position models fairness, it follows that we must extend the original position to nonideal theory. We will now begin completing this task.

§4. The Nonideal Theoretic Original Position – Part II: The Veil of Ignorance and the Parties' Overall Deliberative Situation

We saw in §2 that the parties to a nonideal-theoretic original position must (1) have Rawls' principles and priority relations as an aim; and also (2) be free to weigh those principles and priorities against any and every sort of personal cost s/-he might face under nonideal conditions. In order to provide a clearer picture of how the parties should deliberate given these motivations, we must now further describe their deliberative situation – specifically, their situation behind a nonideal theoretic variant of Rawls' "veil of ignorance".

⁴⁰ Ibid: 19-20.

The nonideal-theoretic veil should in most respects be identical to Rawls' ideal-theoretic veil. No one should have any self-identifying information (such as knowledge of their race, gender, social class, etc.) that might enable them to privilege themselves over others. The only difference between Rawls' ideal-theoretic veil and the nonideal theoretic veil should be the social conditions to which the two are applied. Whereas Rawls' ideal-theoretic parties were assumed to deliberate about well-ordered conditions, the parties to the nonideal theoretic stage should deliberate behind a veil applied to *non*-well-ordered conditions: conditions that fall short of full compliance with Rawls' principles and priority relations. Indeed, the parties should suppose that they could turn out to live under *any* realistically possible nonideal conditions. This ensures that their deliberations are perfectly general. Whatever principle(s) they agree to will be the principles for dealing with *any and every* type of injustice. Does this abstract too far? As we will see in §5, then answer is no. The parties can deliberate to a fully general and plausible, principle of nonideal theory from such an abstract position.

Let us now turn to a puzzle that the combination of the parties' motivations and the veil might appear to pose. We have established that the parties must have Rawls' principles and priority relations as *an* aim, one to weigh against personal costs they might face under nonideal conditions. Now, however, we are imagining them as deliberating about conditions (e.g. a society) in which many people do *not* have Rawls' principles and priority relations as an aim (nonideal conditions are, after all, conditions in which many people seek to *uphold* injustice). Accordingly, it appears that the parties simultaneously have one set of motivations, but must deliberate as though they could turn out to be someone who has the *opposite* motivations. Because no one can simultaneously have and

not have a single motivation (that's a contradiction), the nonideal theoretic original position might appear incoherent. However, this contradiction is only an illusion. There are two possible ways of avoiding the incoherence, one of which is superior to the other. One possibility is to suppose that the parties treat themselves as though they could only turn out to be people who *do* have Rawls' principles as an aim once the veil is raised. In short, they might rule out the possibility that they could turn out to be any person who favors injustice. This "solution", however, does not sit well with a commonsense moral idea. However much we may believe that justice is preferable to injustice, we do not tend to think that people who prefer justice can arbitrarily impose any and every cost they like on people who defend injustice. That is to treat people who prefer justice as something like dictators over the question of how personal costs should be distributed under nonideal conditions. A far more plausible idea is to consider only some interests of unjust people matter: their *just* interests, or the interests they *would* have if they accepted the correct principles of justice. This suggests a better solution. The nonideal-theoretic original position is made coherent by imagining the parties – all of whom have Rawlsian ideals and priorities as an aim – as desiring all of the "unjust" persons they could turn out to be to (1) *come* to accept Rawlsian ideals and priorities, thereby (2) giving up whatever unjust or illegitimate preferences they might have. Indeed, this is a commonsensical idea. Anyone who accepts a particular conception of justice will surely want all "unjust" people to *stop* having unjust ends and instead adopt just ends.⁴¹ The nonideal-theoretic original position, as we have described it, models exactly this. Indeed, there is a simple

⁴¹ In nonideal conditions, the unjust stand in our way. Our task is to pursue their, and our, legitimate interests. We model this by having the parties (a) represent *all* people, and (b) imagining the parties as desiring *every* person they could turn out to be to have their (the parties') just ends.

way to describe the parties' overall deliberative situation behind the veil of ignorance. Because the parties to the nonideal-theoretic original position all have Rawls' principles and priority relations as an end, but deliberate behind a veil of ignorance that makes them assume that they could turn out to be a "just" or "unjust" individual, the parties should have the following general goal for nonideal conditions:

(A) Cause people who do not have Rawls' principles or priorities as an aim to come to have those principles and priorities as an aim.

But this is not the only goal the parties should have. Because, as we saw in §2, every person who has Rawlsian ideals and priorities as aim has a legitimate interest in (a) weighing those ideals and priorities against personal costs, and (b) effectively pursuing their favored such weighting in practice, the parties to the nonideal theoretic original position also have the following two general goals in addition to (A):

(B) Enable all people who have Rawls' principles and priority relations as an aim to weigh those principles and priorities against personal costs; and

(C) Enable all of these same people to successfully pursue their favored weightings.

In summation, the parties want to bring everyone who does not have Rawlsian ideals as an end to come to have those ideals as an end (Goal A), while at the same time enabling people who have Rawlsian ideals as an end to weigh those principles against personal costs (Goal B) and successfully pursue their favored weightings (Goal C). Accordingly, let us describe "nonideal-theoretic primary goods" as all-purpose means for equally advancing goals (A)-(C). The question now is: which sorts of things are nonideal-theoretic primary goods?

§5. Three Nonideal-Theoretic Primary Goods, and Priority Relations Among Them

The question at hand is this: what all-purpose goods generally enable people to (A) cause people to come to have Rawlsian ideals, (B) rationally weigh Rawlsian ideals against personal costs, and (C) successfully pursue whatever weighting of Rawlsian ideals against personal costs one favors? One tempting answer might be Rawls' *own* primary goods: basic rights and liberties, opportunities, income and wealth, and social props to self-respect. After all, suppose an oppressed person with Rawlsian ideals comes to have more rights and liberties, opportunities, or wealth than s-/he previously had. Won't s-/he then be in a better overall position to promote Rawlsian ideals, as well as whichever weighting of Rawlsian ideals against personal costs s-/he most favors? We have already seen that this is not the case. For again, as we saw in §2, it all depends on how the person comes to *get* those goods. If the manner in which the person comes to enjoy Rawlsian primary goods imposes personal costs on the person that the person prefers not to endure, then the mere act of giving that person Rawlsian primary goods will undermine that person's legitimate goals (i.e. his/her weighting of Rawlsian ideals against personal costs). Therefore, the parties to a nonideal-theoretic original position should not simply distribute Rawlsian primary goods. For all they know behind the veil of ignorance, the distribution of Rawlsian primary goods may undermine their own goals once the veil is raised.

The problem we now face is that it looks as though nothing could qualify as a nonideal-theoretic primary good. After all, it seems as though there are potential personal costs associated with a person coming to enjoy just about anything worthwhile. For example, someone could presumably want to come to enjoy equal political rights but

oppose certain kinds of tactics (e.g. intimidation, terror) for obtaining those rights. The simple fact of the matter is that people often have strong, even impassioned, preferences about *how* (ideal) justice is achieved. Some are willing to endure great sacrifices for the cause of justice (e.g. being arrested or beaten by police); others are not so willing. Some want progress to occur slowly, so that important traditions and personal relationships are not destroyed⁴²; others care very little for tradition or relationships. And so on. The kinds of costs people are willing to endure for social progress vary immensely from person to person.

The parties to the nonideal-theoretic original position may not be able to entirely eliminate the possibility of cutting their own goals once the veil of ignorance is raised. They may, however, be able to *minimize* the possibility. How so? At any particular point in time, a person with Rawlsian ideals will be best able to simultaneously (i) promote Rawlsian ideals while (ii) advancing his/her favored weighting of those ideals against personal costs by (iii) *working together with other people who have the same ideals and favored weightings that they do*. For instance, suppose I am a woman in a patriarchal society who wants women to come to enjoy the same rights as men (in conformity with Rawlsian ideals), but who favors facing some particular set of personal costs C for the sake of realizing those ideals over all other types of costs. Regardless of what C turns out to be, I will be *better* able to ensure that processes of social change imposes only C on me (rather than any other set of less-preferred costs) the more (a) similarly-minded people I

⁴² See www.we-change.org/spip.php?article326, for a moving personal account from a female women's rights activist in Iran who believes it is of great importance for her to maintain a positive relationship with her (chauvinist) father in the course of social progress. Although there will of course be some question about the moral legitimacy of putting these sorts of personal issues "in front of" the cause of women's liberation, we have, at least in Rawlsian theory, as we have seen in §2 of this paper, no reason to consider such concerns to be illegitimate (see §2 of this paper).

have on my side, and the more (b) social progress conforms to our collective decisions as a group. There is, we say, strength in numbers.

This brings us to something that is intuitively morally important about the idea of grass-roots social movements: specifically, those that are broadly *equitably* responsive to their members. When a social movement is driven from the ground up, by the interests or preferences of the many rather than by decisions “on high” by a few, that movement is more likely to be *sensitive to the costs that each of its members are willing to endure for the movement’s ends, or goals*. The parties to the nonideal-theoretic original position should then to have sufficient reason to regard the following as genuine nonideal-theoretic primary goods: (a) opportunities to participate *equitably*⁴³ *and effectively* in a broadly equitable grass-roots Rawlsian social movement, and (b) conformity of social progress to their favored group’s collective decisions. Although participation in a grass-roots social movement may itself impose significant personal costs on people, a person can always choose to abstain from participating in a social movement (by simply “staying home”). Again, the point is *not* that opportunities to participate in such groups totally eliminate the problem of people facing personal costs they disfavor. The point is that (a) having people on one’s side (i.e. being a member of an effective but equitable grass-organization dedicated to Rawlsian ideals), and, (b) having social progress conform to the collective decisions of those on one’s side (i.e. conformity to the group’s collective decisions), appear to be the two most efficacious ways for *any* person to shape the course of social progress in a manner that conforms to their favored weighting of Rawlsian

⁴³ Equitably because, behind the veil of ignorance, the parties know they could be “anyone” once the veil is raised – in which case they should want *every* person they could turn out to be to have a real “say” over the course of social progress.

ideals against personal costs. Since these two goods qualify as nonideal-theoretic primary goods, the parties to the nonideal-theoretic original position should aim to bestow them on every person they could turn out to be.

It is important to note here that the parties have sufficient reason to prioritize the first of these two goods (the opportunity to participate in an effective but equitable Rawlsian group) over the latter (conformity to a group's collective decisions). After all, the parties should not want social progress to conform to the collective decisions of an *inequitable* group. Given that the parties could be *anyone* once the veil of ignorance is raised, the parties should want to ensure that *every* person they could turn out to be is taken seriously and has an opportunity to have a "say" over the costs they will endure for social progress. Thus, from the perspective of the nonideal-theoretic original position, opportunities to participate *effectively and equitably* in a Rawlsian reform group should take lexical priority over the good of having social progress conform to the decisions of such a group.

We will now see that a third nonideal-theoretic primary good exists – one that takes intermediate priority between these two just discussed. In addition to opportunities to participate equitably and effectively in (Rawlsian) grass-roots groups, the parties should want each group to have a few determinate common aims – aims involving education. The parties should want such groups to educate people inside and outside the group in (a) Rawlsian ideals, (b) skills and information useful for rationally weighing Rawlsian ideals against personal costs and (c) skills and information useful for pursuing such weightings effectively. Consider the example of a woman in a patriarchal society who is a member of a (broadly) Rawlsian grass-roots reform movement. The group she

belongs to seeks equal rights for women. However, many women in her society claim to *not* to want equal rights. Why? One possibility is that many of these women are under-educated and have no conception of what life would be like in a society in which men and women are treated equally, both under the law and in custom. Now imagine that the women who claim to not want equal rights (many of whom are under-educated) became educated in what an equitable society would look like, and understood that they had a choice shaping the process that brought about such a society—including the right to decide what types of sacrifices they would be willing to bear. Perhaps many of them would become more open to receiving equal rights. At the very least, they would have enough knowledge to make an informed decision about their views on obtaining equal rights. The point of the example is this: because the parties to the nonideal-theoretic original position know they could turn out to be anyone once the veil of ignorance is raised (even these women, who reject Rawlsian ideals), the parties should want Rawlsian grass-roots groups to have the *education* of their members and their society at large as a common, fundamental aim. This education consists of: (a) education in Rawlsian ideals, but also (b) in skills and information for weighing Rawlsian ideals against personal costs (and effectively pursuing whichever weighting a person favors). In practice, these types of education might comprise pamphleteering on women's rights, seminars on economic and social policy, and so on. As I stated earlier, these educational aims are to take lexical priority over the nonideal-theoretic primary good of having social progress conform to the group's collective decisions. This is for an obvious reason: if the members of a (Rawlsian) grass-roots movement are not adequately educated on (a) Rawlsian ideals, (b) skills and information relevant to rationally weighing Rawlsian ideals against personal

costs, and (c) skills and information relevant to pursuing their favored weightings effectively, there is a real moral question of whether the members of the group are *competent*. The parties to the nonideal-theoretic original position do not want social progress to conform to the collective decisions of a group whose members are inadequately educated in relevant matters. Furthermore, if this education extends beyond the members of the group to society at large, it follows that many (formerly) under-educated individuals would have adequate knowledge to decide whether or not they supported a particular group's movement and aims. Indeed, if the general population of a society was provided an adequate education on the areas listed above, it is plausible to think that at least some (if not many) of the previously under-educated and dissident population would join the ranks of the Rawlsian grass-roots group. This, in turn, would strengthen the likelihood that social progress would conform to the collective decisions of the group. Education on those matters, then, is a nonideal-theoretic primary good which is *intermediate in priority* between (a) opportunities to participate equitably and effectively in a Rawlsian grass-roots reform group, and (b) having social progress conform to the collective decisions of such a group.

In summation, we have established three nonideal-theoretic primary goods and priority relations between them. These goods are: (i) opportunities to participate equitably and effectively in a Rawlsian grass-roots reform group; (ii) such a group having education in Rawlsian ideals, as well as skills and information relevant to rationally weighing Rawlsian ideals against personal costs, as a *common aim* (i.e., for themselves and for society at large); and (iii) having social progress conform to the collective decisions of a Rawlsian group that satisfies these conditions (i.e. a group that is not only

equitable and effective, but has the aforementioned educational aims as a common end). The final question is how the parties to a nonideal-theoretic original position should want these goods distributed across society. We now turn to this final issue.

§6. Three Arguments for the General Principle of Rawlsian Nonideal Theory

The final question for us is whether the parties to the nonideal-theoretic original position could come to an agreement on some determinate principle(s) for distributing the nonideal-theoretic primary goods just defended. I will now argue that there are two arguments to be made for the following principle and, by extension, the three corollaries that follow:

General Principle of Rawlsian Nonideal Theory: For any class of unjust conditions U , the more a given population P is unjustly disadvantaged under U , the more goods for (A) causing others to come to have Rawlsian ideals, (B) enabling those who have Rawlsian ideals to weigh those ideals against personal costs, and (C) enabling those who weigh Rawlsian ideals against personal costs to successfully pursue their favored weightings, the members of P ought to enjoy.

First Corollary (The Grass-Roots Principle): The more a given population P is unjustly disadvantaged under nonideal conditions U , the more opportunities the members of P ought to have to participate *equitably and effectively in a (broadly) Rawlsian grass-roots social movement*.

Second Corollary (The Education Principle): The more a given population P is unjustly disadvantaged under nonideal conditions U, the more the members of P ought to have the common aim of *educating* people in: (a) Rawlsian ideals, (b) skills and information relevant to rationally weighing Rawlsian ideals against personal costs, and (c) skills and information relevant to effectively pursuing such weightings.

Third Corollary (The Grass-Roots-Decision Principle): The more a given population P is unjustly disadvantaged under U, the more the overall course of social reform ought to conform to the *collective decisions* of grass-roots movements that satisfy the first Two Corollaries.

§6.1. The Uniquely Powerful Case for “Maximin” Within Nonideal Theory

In §26 of *A Theory of Justice* – the section entitled, “The Reasoning Leading to the Two Principles of Justice – Rawls famously argues that, “it is useful as a heuristic device to think of [his] two principles as the maximin solution to the problem of social justice.”⁴⁴ Indeed, although there is increasing consensus that Rawls does not in the end hang his argument on maximin⁴⁵, Rawls does spend §26 of *A Theory of Justice* arguing that *maximin* (the rule of pursuing the most superior of all possible *worst* outcomes) is the most rational rule for the parties to the original position to adopt. Rawls writes,

⁴⁴ Rawls (1999a): 132.

⁴⁵ Again, see the Stanford Encyclopedia of Philosophy entry, “Rawls”, §4.7.

[T]here appear to be three chief features of situations that give plausibility to this unusual rule. First, since the rule takes no account of the likelihoods of the possible circumstances, there must be some reason for sharply discounting estimates of these probabilities....The second feature that suggests the maximin rule is the following: the person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule...the third feature...[is]...that the rejected alternatives have outcomes that one can hardly accept. The situation involves grave risks.⁴⁶

Rawls then argues that the deliberative situation of the parties to the original position, within ideal theory, satisfies all three of these conditions “to a very high degree,”⁴⁷ and that his two principles of justice and their priority relations are the outcome of this type of reasoning.⁴⁸

Now, of course, Rawls’ complete case for his principles of justice and their priority relations is quite complex. Rawls actually gives several different arguments for his principles: one from “strains of commitment,”⁴⁹ one from considerations of “publicity,”⁵⁰ one from considerations regarding “self-respect,”⁵¹ and one from a conception of citizens’ “needs.”⁵² We have neither time nor need to review each of these arguments here. However, it is worthwhile to briefly raise a couple of trenchant worries

⁴⁶ Rawls (1999a): 134.

⁴⁷ Ibid: 133.

⁴⁸ See especially *ibid*: §§33-5.

⁴⁹ See *ibid*: 153-4.

⁵⁰ See *ibid*: 154-5.

⁵¹ See *ibid*: 155-6 and 477-79.

⁵² See Rawls (1993): 187-90.

about Rawls' arguments in ideal theory, so that we can see whether the same objections hold within nonideal theory.

One worry is that the first condition that Rawls adduces in favor of maximin – the claim that there are reasons for sharply discounting probabilities of various outcomes – is not satisfied by the original position. Although Rawls argues against the idea that behind the veil of ignorance, it is rational for the parties to assign an equal probability to every possible result (in accordance with the so-called “principle of insufficient reason”), several critics have argued it *is* rational for the parties to reason in such a way.⁵³ Secondly, critics argue that the original position doesn't satisfy the second *or* third conditions that Rawls defends, at least with respect to the primary goods of income and wealth.⁵⁴ For consider again what the second and third conditions are. Rawls claims that maximin is rational when, in addition to discounting probabilities, one cares *little* about doing better than whatever result maximin assures and one cares very *much* about avoiding results worse than that which maximin guarantees. As many critics have pointed out, when it comes to economic matters, there seems to be little reason to think that either of these things is true of the parties to the original position. On the one hand, many people the parties could turn out to be once the veil is raised presumably *would* want to do better than maximin might allow (I might want to be very, very rich, for example, more rich than Rawls' difference principle might allow). Similarly, on the

⁵³ See Brian Barry, *The Liberal Theory of Justice: A Critical Examination of the Principal Doctrines in "A Theory of Justice" by John Rawls* (Oxford: Clarendon Press, 1973): chapter 9; and John Harsanyi, "Can the Maximin Principle Serve as a Basis for Morality? a Critique of John Rawls' Theory," *American Political Science Review* 69, 2 (June 1975): 594-606.

⁵⁴ See Will Kymlicka, *Contemporary Political Philosophy* (Oxford: Oxford University Press): 66 for a brief summary of these worries. For the original criticisms, see Barry (1973): ch. 9, and R.M. Hare, "Rawls' Theory of Justice", in N. Daniels (ed.), *Reading Rawls* (New York: Basic Books, 1975): 88-107.

other hand, it is not clear why anything less than maximin (i.e. Rawls' difference principle) would be intolerable for the parties. Why, many critics ask, wouldn't some social minimum *less* than maximin be tolerable (for the parties)? Starvation and homelessness, for example, seem genuinely intolerable – yet Rawls' difference principle requires far more than these sorts of social protections: it requires *maximizing* the wealth of the least well-off (in perpetuity). It is hard, for many of Rawls' critics at least, to see why the parties should regard anything less than the difference principle to be intolerable.

Rawls' case for maximin within ideal theory, then, lends itself to many objections. But what about *nonideal* theory? Do the same objections apply? Let us begin with the third of Rawls' three conditions and work our way backwards: the question of whether “sub-maximin” results (i.e. payoffs they could face as a result of choosing some other rule) should be regarded by the parties as genuinely intolerable. We do not need to assume that Rawls' arguments within ideal theory are actually successful to make a compelling case for this condition within nonideal theory. Consider that whatever injustice turns out to be (we may accept Rawls' principles or we may not), injustice *by definition* is something that the parties to the nonideal-theoretic original position cannot tolerate. For the parties are, once again, *motivated* by principles of ideal theory (within our specific framework, Rawls' principles and priority relations). Moreover, the worse a given injustice is, the *more* intolerable the parties should regard those conditions (for again, by definition, the more unjust conditions are, the *more* those conditions fall short of the parties' own ideals). Rawls' third condition – call it the Intolerability Condition – thus seems satisfied in the nonideal-theoretic case.

Now turn to the second condition that Rawls discusses: the question of whether (from the perspective of the nonideal-theoretic original position) the parties have reasons to not care much about doing better than maximin guarantees. In order to evaluate this condition, consider briefly the situation the parties to the nonideal-theoretic original position face: they know that they could turn out to be anyone under nonideal conditions (and that every person they turn out to be should be motivated by Rawlsian ideals). Every such person they could turn out to be, then, is either (a) unjustly advantaged by injustice in some way, or else (b) unjustly disadvantaged in some way. Now, because they are motivated by Rawls' principles and priorities as an end, the parties should clearly sharply discount the interests (or preferences) of all persons who are unjustly advantaged under nonideal conditions. From their (the parties') point-of-view, these people already enjoy advantages they shouldn't have. The real question then is how to rank outcomes for the unjustly *disadvantaged*. Now, I submit that there are clear reasons for the parties to adopt "the law of diminishing returns" for people who suffer injustice. Let me explain. It is clear, offhand, that the more a person is disadvantaged by injustice, the *more* they have to gain (personally speaking) from social reform. A person who is positively oppressed, for example – a slave, for example – surely gains more from being emancipated than, say, the unjustly discriminated against job applicant would gain from no longer being discriminated against. Both are clearly important gains – but clearly, the former gain is far greater than the latter. As a general rule, the more a person is oppressed, the more they have to gain by being empowered to stand up for their just interests with other like-minded people (through the three nonideal-theoretic primary

goods defended here). Thus, I submit that Rawls' second condition in favor of maximin (and so, in favor of my General Principle) is well satisfied within nonideal theory.

Finally, there is Rawls' first condition: the supposition that there are compelling reasons to sharply discount probabilities. This condition is also satisfied in the nonideal-theoretic case. Since one does not know precisely which sorts of nonideal conditions one will live in – we did, again, abstract away from particular nonideal conditions and instead have the parties deliberate on the assumption that they could turn out to live under *any* nonideal conditions – the parties have no reason (at all) to assign a particular probability to outcomes. Moreover, it could be positively disastrous for them to adopt the principle of insufficient reason (the assumption that all outcomes have the same probability) since, as we have just seen, some outcomes (outcomes for those who are the *most* disadvantaged by injustice) are far more important to the parties than other outcomes (e.g., outcomes for the unjustly *advantaged* and those who are *less* unjustly disadvantaged than the worst off).

I propose, therefore, that as open to objections as Rawls' case for maximin may be within ideal theory, the case for maximin is wonderfully powerful within nonideal theory. If this is right – and I believe it is – the parties to the nonideal-theoretic original position ought to agree to the General Principle of Rawlsian Nonideal Theory and its three corollaries. For the General Principle simply is *maximin* as applied to nonideal-theoretic primary goods, and the three corollaries follow from the General Principle by the priority relations among nonideal-theoretic primary goods defended in §5 of this paper.

§6.2. The General Principle and Corollaries in Reflective Equilibrium

The General Principle of Rawlsian Nonideal Theory coheres with a number of pre-theoretic moral intuitions about the idea of “just social progress.” It coheres, first, with the simple fact that we *celebrate* grass-roots social movements (for the correct ideals). Consider, for example, the U.S. Civil Rights Movement in the 1960s. Although it was a monumental struggle, one that involved many costs – police abuse, harassment, lynchings, etc. – we celebrate this movement and others like it across the world and across history (the US Revolutionary War is another case) because we think there is something right and just about *the oppressed liberating themselves*. Contrast this to cases where people are “given” justice by (e.g. the force of) others. “Operation Iraqi Freedom”, the United States’ 2003 invasion and subsequent occupation of Iraq, is one such case. Why are so many of us so skeptical about the justice of the Iraq invasion and occupation? One obvious answer is this: it was a decision to the Iraqi people themselves – the oppressed – played no part. It would have been one thing had the Iraqi people staged a revolution against Saddam Hussein themselves. I think we would all agree that that would have been just (notice that the General Principle I defend confirms this). It also would have been one thing if the Iraqi people had *requested* international intervention (as might have occurred if leaders of an underground Iraqi movement had repeatedly petitioned the international community for help, much as the African National Congress during Apartheid in South Africa⁵⁵). In that case, I think many of us would be tempted to say that an invasion would be just (and notice: this is also confirmed by the

⁵⁵ Leaders of the ANC repeatedly addressed the UN General Assembly during the 1960s, requesting (among other things) an international embargo on South Africa.

General Principle). But of course none of this occurred. The US invaded Iraq of its *own* accord, with little regard to what anyone (let alone Iraqis themselves) thought about the matter. This is, clearly, an enormous reason why so many people condemn the US's actions. Social progress should not be forced on people "on high." It should be driven by, or responsive to, the needs of the oppressed themselves – just as the General Principle requires.

Of course, we do not always celebrate grass-roots movement. We have significant moral worries whenever (a) a "grass-roots" movement is dominated by a few powerful members who apparently care little about what the members at large seem to think, or whenever (b) the members of the group seem to base their views and preferences on false or inadequate information. Yet the General Principle and Three Corollaries still coheres with these points. The General Principle and Three Corollaries tell us that reformist groups should be equitable and have various forms of education as a primary common aim. I submit, as such, that the General Principle and Three Corollaries, quite aside from following from the nonideal-theoretic original position by way of a compelling maximin argument, cohere with many convictions we have about nonideal justice.

Conclusion

We have seen here that Rawlsian nonideal theory has been missing a crucial component – an analysis of transitional fairness – and how to amend this gap. First, we discussed why the original position must be extended to nonideal theory. We then extended the original position to nonideal theory, showing how its parties ought to prioritize a class of nonideal

theoretic primary goods above Rawls' principles and priority relations. Next, we saw that the parties should pursue three concrete nonideal-theoretic primary goods, as well as priority relations among them. We then saw that whatever problems there are with Rawls' arguments in ideal theory, a very strong case can be made for *maximin* within nonideal theory (and so, for the General Principle and Three Corollaries I defend). Finally, we saw that the General Principle and Corollaries fare well in reflective equilibrium. I propose, as such, that we have, at long last, foundations for a plausible, and illuminating, Rawlsian nonideal theory of justice.