

Rectifying Wrongs: The Problem of Historical Injustice

Moises Vaca Paniagua

*PhD Thesis
University College London*

I, Moises Vaca Paniagua, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis is concerned with the problem of rectification in the theory of justice. We are faced with examples of great historical injustice over the last few centuries. A proper regard for the demands of rectification seems required of us in the face of the overwhelming importance that victims place on it; without it, no society can hope to sustain mutual respect among its citizens, the non-victims and the victims, nor probably foster the self-respect of the victims. I argue that the problem of rectification poses a distinctive and fundamental problem for classical theories of justice and specifically for John Rawls's account of justice-as-fairness. Defenders of Rawls might claim, first, that rectification falls outside the scope of his theory of justice, since that is intended as ideal theory, and thus formulated against the fictional assumption that no historical wrongs have taken place. In this view, rectification is a concern of real political theory but not of ideal theory of justice. I argue that this defence is mistaken. Secondly, defenders of Rawls who concede that rectification is a proper part of the ideal theory of justice might claim that the principles of justice-as-fairness provide a basis for

determining the extent to which justice requires rectification of wrongs. This too, I argue, is mistaken. In light of the demands that rectification places on us, I propose an alternative picture of equality as conceived of within the liberal tradition.

Contents

ACKNOWLEDGEMENTS	7
ABBREVIATIONS FOR RAWLS'S BOOKS	10
INTRODUCTION	11

PART I HISTORICAL RECTIFICATION AND LIBERAL THEORY

1 THE NORMATIVE SIGNIFICANCE OF HISTORICAL INJUSTICE	24
1.1 Victims' Psychology and Self-Respect	24
1.2 Restoring Mutual Respect and <i>Recent-Past Rectification</i>	40
1.3 Liberal Integrity and <i>Critical Historical Narratives</i>	49
1.4 Concluding Remarks	60
2 IS HISTORICAL INJUSTICE OUTSIDE THE SCOPE OF RAWLS'S THEORY?	61
2.1 The <i>Ideal Character</i> of Rawls's Theory	62
2.2 The <i>Forward-Looking Character</i> of Rawls's Theory	84
2.3 The <i>Non-Comprehensive Character</i> of Rawls's Theory	98
2.4 Concluding Remarks	107

3 TOWARDS A PRINCIPLE OF HISTORICAL RECTIFICATION FOR IDEAL THEORY	108
3.1 The Two Principles and Historical Rectification	109
3.2 Towards a Principle of Historical Rectification for Ideal Theory	119
3.3 On a Recurrent Transitional (non-Ideal) Problem: <i>Retribution VS Recognition</i>	125
3.4 Concluding Remarks	134

PART II EQUALITY AND PRACTICE-DEPENDENCE

4 THE PROTECTION OF CITIZENS' MORAL EQUALITY	137
4.1 The <i>No-Use Critique</i> and the Fundamental Egalitarian Aim	138
4.2 The <i>Institutional Injustice Critique</i> and the Two Principles	147
4.3 The <i>Non-Institutional Injustice Critique</i> and the Egalitarian Ethos	157
4.4 Concluding Remarks	165
5 PRACTICE-DEPENDENCE, CONSTRUCTIVISM, AND LIBERALISM	167
5.1 A Concept of <i>Pure Justice VS</i> a Concept of <i>Impure Justice</i>	168
5.2 A <i>Constitutive VS</i> an <i>Indicative</i> Account of Correctness	180
5.3 A <i>Comprehensive VS</i> a <i>Political</i> Conception of Justice	191
5.4 Concluding Remarks	203
REFERENCES	204

Acknowledgements

This thesis owes very much to the guidance and support of my primary and secondary supervisors, Veronique Munoz-Dardé and Michael Martin. I have learned from Veronique Munoz-Dardé ever since the first conversation we had the day I arrived at UCL. I can only hope that this work reflects her thoughtful advice over the last couple of years. I would also like to thank her for her fundamental role in facilitating my academic stays at the Australian National University, the University of Michigan, and the University of California (Berkeley), as well as for her endless institutional support in various other moments.

Thanks to Michael Martin for the insightful year we spent working on mind, metaethics and reasons theory, and for making me see why we should have done so in that order and not the reverse, as I originally wanted. Likewise, I owe a great debt to the professors with whom I worked at UCL prior to my thesis: Michael Otsuka, Marcus Giaquinto, and Mark Kalderon. Thanks also to Jonathan Wolff, Paul Snowdon, José Zalabardo, and Rory Madden for their incredible helpfulness.

My special gratitude goes to the supervisors of my previous two degrees: Faviola Rivera and Will Kymlicka. Faviola Rivera not only set me on the right track as an undergraduate with interests on ethics and political philosophy at UNAM but has ever since continuously helped me far beyond the call of duty. I will always be in debt to her for this. Both the work and kindness of Will Kymlicka, with whom I worked as a master's student at Queen's University, have fundamentally shaped my understanding of how to do political philosophy and be a responsible academic.

Fellow students at UCL and Oxford were also very important in my formation. I learned a great deal from seminars, conferences, reading groups, and discussions with Laura Valentini, Chris Jay, George Hull, David Holly, Jeremy Farris, and Robert Jubb. To my dear friends Marion Godman, Paloma Atencia-Linares, Mario Santos-Sousa, and Gabriel Wollner, I owe special thanks. Aside from having made my life in London intellectually engaging, they also made it much nicer.

From the Institute of Philosophical Research at UNAM, I thank Carlos Pereda and the members of the junior seminar: researchers Eduardo García-Ramírez, Ángeles Eraña and Axel Barceló. They, together with Miguel Ángel Fernández, were instrumental in many presentations at the Institute regarding work in this thesis. To my dear friends Eduardo García-Ramírez and Catalina Pereda I owe much more ever since we started as undergrads at UNAM. I also deeply thank my friend Juan Espíndola: a long and revealing conversation in Ann Arbor prompted us to join forces in a single paper. That

paper triggered the philosophical exercise that I have developed in these pages.

Finally, I must mention my parents, Gloria and Sergio, as well as my brother and sister-in-law, Felipe and Irasema, for their astonishing encouragement and support. I would simply not have done a PhD without their help. My partner, Mara, was right by my side on the first day I started working on this thesis and remained there until the very last one. Now I can only hope to do the same myself for her. To these five people, my core people, I dedicate this thesis.

I would like to state that my PhD studies were fully funded by a scholarship (205636) from the National Council of Science and Technology of Mexico, CONACYT. I have also received support from the UNAM projects PAPIIT 402709: 'Liberalism in Mexico: history, theory and contemporary problems'; PAPIIT IN401611: 'Representation and cognition'; and PAPIIT IA400112: 'A cognitive approach to philosophy of mind'.

Abbreviations for Rawls's Books

- CP* *Collective Papers*. Harvard University Press, 1999.
- JF* *Justice as Fairness. A Restatement*. Harvard University Press, 2001.
- LP* *The Law of Peoples*. Harvard University Press, 1999.
- PL* *Political Liberalism*. Columbia University Press, 2005.
- TJ* *A Theory of Justice*. Revised Edition. Harvard University Press, 1999.

Introduction

Most contemporary liberal societies are rife with civic divisions related to the past: there are always groups of citizens who in the past have been victims of some gross forms of distributive injustice; citizens, that is, who have been affected by flagrant violations of the liberal principles that are to regulate social cooperation. Among these injustices, particularly brutal are those characterised as forms of *political violence* or *political terror*. What makes these injustices of a similar kind is that they are perpetrated and directed either by the State (through one or more of its institutions) or by groups of citizens (with the institutional support or consent of the State) towards another particular and identifiable group of citizens. These injustices may vary in their targeted groups and gravity: first, they can involve the institutional promotion of the exploitation or extermination of a cultural, religious, or ethnic group; second, they can involve the persecution, imprisonment or abduction of citizens because of their political beliefs and actions.

Examples of the first type of injustice include the exploitation of indigenous peoples by most Latin American states in the 19th century and the extermination of Mayan persons by the Guatemalan State from 1970 up to 2000. Examples of the second type of injustice include the imprisonment and abduction of individuals considered to be political dissidents or “subversive” elements by most Latin American states (especially, Argentina, Chile and Uruguay) in the 1970s and 1980s.

An immense amount of evidence in psychology and the social sciences, as well as in moral and political philosophy bolsters the idea that coming to terms with a past shaped by political violence is a phenomenon of the utmost importance for citizens and liberal societies. Most of the work on this phenomenon falls under the name of transitional, restorative, or compensatory justice. It seems to be an assumption of the debate that, however widespread and deep this phenomenon is, it falls out of the competence of the most prominent theory of liberal justice, i.e. Rawls’s theory. I think this is a mistake.

In fact, I have come to believe that it is not immediately obvious that the full regulation of Rawls’s justice-as-fairness in a given society would secure the rectification of important historical injustices such as the ones just mentioned. Since, as I will argue, the scope of justice needs to encompass historical rectification, Rawls’s description of a well-ordered society might be depicting an incomplete ideal of justice to which we should try to arrive at in our actual social proceedings. If this claim is correct, then Rawls’s theory

must show concern for historical injustice. This is the main claim that I defend throughout this work.

To that end, in CHAPTER 1 I first present the empirical and normative evidence granting the significance of historical rectification. In particular, I claim that coming to terms with a past shaped by political violence bears an extremely high importance to citizens and liberal societies in relation to (1) citizens' psychology and self-respect, (2) citizens' mutual respect, and (3) society's liberal integrity as well as the achievement of social stability for the right reasons. You may find this way of proceeding to be somewhat odd for a philosophical text —i.e., paying attention to the empirical evidence provided by psychological studies. However, sensitivity to psychological and social facts is one of the major features distinguishing Rawls's conception of justice from others. In fact, such sensitivity is now regarded as a shared assumption and a cherished characteristic of the approach to political philosophy undertaken by all the so-called *practice-based* or *practice-dependent* accounts of justice. My attempt to inform our normative assessment of historical rectification through psychology and not only through more independent philosophical considerations aligns with this way of doing political philosophy.

Having defended the normative significance of past rectification, in CHAPTER 2 I revise a first Rawlsian objection to my main claim (i.e., that Rawls's theory is to show concern for historical injustice). This objection — generally the default position in contemporary political philosophy— states

that, regardless of the importance of historical injustice in the real world, its rectification is a problem that does not arise within Rawls's theory.

Three considerations may be given to support this objection. These considerations appeal, respectively, to the *ideal*, the *forward-looking*, and the *non-comprehensive* character of Rawls's theory. In particular, the *ideal* character of Rawls's theory has a powerful effect on both sides of the debate on whether or not past wrongs hold normative significance for present justice. Like their opponents, a vast majority of authors defending the normative significance of past wrongs for present justice hold that, in order to address such preoccupations, we need to drop the assumptions of Rawls's *ideal theory* all together. Consequently, it is presumed that the work on historical injustice falls exclusively under the realm of *non-ideal theory* and, thus, is outside of the scope of Rawls's theory.

Against this objection I argue that the *ideal*, *forward-looking*, and *non-comprehensive* character of Rawls's theory does not justify its lack of concern for historical injustice. There are two main reasons for this. First, victims' interest in coming to terms with a past shaped by political violence is on a par with other human psychological regularities that are already central to Rawls's *ideal theory*. Thus, I hold that such a theory should consider the psychological regularities associated with the importance of coming to terms with an unjust past just as much as it considers other psychological regularities. Second, I argue that an *unjust pre-ordered past* is a permanent feature of a liberal well-ordered society, in the very same way that *reasonable*

and *cultural pluralism* are. Considering these two reasons, I hold that we as theorists need to show that even under historical conditions marked by gross injustice —as is the case in all current liberal societies— the Rawlsian utopia of a stable well-ordered society is still realistic. Accordingly, the conclusion of this chapter is that a correct ideal non-comprehensive theory of liberal justice cannot be exclusively forward-looking.

Having defended that the first objection fails, in CHAPTER 3 I consider a second Rawlsian objection: the idea that, even when historical injustice *is* within the scope of Rawls's theory, the effective regulation of the very principles of justice-as-fairness suffice to ensure its rectification in a given liberal society. For this second objection to succeed, it must be the case that a society in which relevant historical injustices remain un-rectified does not fully realise at least one of Rawls's principles. An obvious candidate in this respect is the principle of fair equality of opportunity. This is so because a non-rectified historical wrong may have tremendous effects on victims' capacity to pursue their plan of life. These effects can introduce unfairness in the competition for social positions of advantage, affecting victims' prospects (and their descendants) but not those of the rest of the citizenry. Thus, the main idea behind this second objection is that the principle of fair equality of opportunity would not effectively regulate the basic structure of society until the State had rectified all relevant historical injustices.

Against conventional assumptions in current political philosophy, we will see how this second objection is more promising than the first one. If this

work manages to persuade you this much, it will have already made a contribution (however modest) to contemporary discussion amongst Rawlsians. For such a discussion should not gravitate around the question of which aspect of Rawls's theory justifies excluding concern for past injustice. Rather, it should centre on the question of whether Rawls's theory as it stands right now contains the normative apparatus to response to them or not.

As promising as the second objection first appears to be, however, I argue that it is not immediately clear that it succeeds. For while it is certainly the case that relevant past wrongs have effects on the opportunity of victims to compete for positions of advantage in the present, this problem is not the only normatively salient consequence of a non-rectified historical injustice. While such injustices affect (1) citizens' psychology and self-respect, (2) citizens' mutual respect, and (3) society's liberal integrity, the second objection incorrectly assumes that the importance of rectifying past wrongs is exhausted by the fact that citizens might otherwise compete under unfair circumstances. Furthermore, I argue that some mechanisms of historical rectification do not seem to follow directly neither from the correct regulation of the fair equality of opportunity principle nor from the correct regulation of the first principle of justice.

Considering this, I explore in a more tentative fashion —also in CHAPTER 3— one way in which Rawls's theory could respond to the challenge posed by historical injustice: adding a specific *backward-looking* principle of

historical rectification to justice-as-fairness. This addition could clarify the commitment of Rawls's theory to the ideal that *no liberal society is fully just until historical rectification has taken place*, so that 'closing the books' or drawing a 'thick line' through the past is never an option from the point of view of justice. The principle could command that recent historical wrongs are to be rectified by the institutional provision of *retribution, compensation*, and *recognition*-driven measures as appropriate. Also, it could establish that the institutions of a well-ordered liberal society promote what I call a *critical historical narrative* (that is, a particular understanding of the past which not only celebrates the putative remarkable historical facts of the society in question but, also, one that acknowledges the past wrongdoing of such a society as a cultural and political continuum over time).

These three chapters form the first part of this thesis. The second part includes two further chapters. In CHAPTER 4 I focus on a similar critique that is often made against Rawls's theory: i.e., its lack of concern for racial and gender relations. I explore three different ways in which this critique can be understood and conclude that one of them prompts the inclusion of an *egalitarian ethos* within the description of a well-ordered society. Importantly, I defend that Rawls's theory already contains all the normative elements required for establishing that such an ethos is informed by a norm commanding respect for citizens' moral equality. In this chapter I also compare my claim that justice-as-fairness should be complemented with a new backward-looking principle with a similar claim made by Seana Shiffrin:

to wit, that a specific anti-discrimination principle with forward-looking character should be included within such a conception.

Finally, in CHAPTER 5 I defend the general *practice-dependent* approach developed in this thesis from G. A. Cohen's recent and powerful criticisms. As I have already mentioned, I support the idea of informing political philosophy in general —and liberal theories of distributive justice in particular— with the evidence and findings of psychology and the social sciences. In this work I press Rawls's theory regarding the importance of historical rectification partly on this basis. For this reason, a defence against Cohen's attack of *practice-dependent* conceptions of justice is in order. In particular, I defend three points: first, that Cohen is mistaken in presenting his critique of Rawls's concept of justice as a critique of Rawls's constructivism. Second, that Cohen's first condition of fundamental principles of justice (stating that such principles must be fact-insensitive) is trivial. And, third, that Cohen's second condition of fundamental principles of justice (stating that such principles must be derived from considerations of justice alone) creates what I call *Cohen's dilemma*. I also defend in this final chapter that a proper political conception of justice should not endorse Rawls's political constructivism neither as a view of the *justification* of principles of justice nor as a metaethical view about what *constitutes* correct principles of justice. For only a deflationist account of the original position —i.e., an account in which the original position is understood as a mere heuristic device of argumentation— is compatible with the aims of political liberalism.

This is the work that you will find in this thesis. Now let me make three caveats to clarify what you will not find in it. First, notice that when I speak of historical injustice, I only refer to non-rectified cases of *past political violence*. I do this for the following reason. Political violence constitutes a gross and clear violation to the precepts of liberal justice. Contrary to regular violence, political violence is always targeted: a person is victim of it either *because* of her belonging to an ethnic, cultural, or religious group, or *because* of her political beliefs and actions. Thus, the very purpose of political violence (again, contrary to regular violence) is to undermine the equal moral status of persons.

Importantly, defending a framework of past rectification focused on these cases does not commit oneself to a particular theory of property or entitlement —say, *a la* Robert Nozick. As it is well known, for Nozick any transaction violating either what he calls the principle of justice in acquisition or the principle of justice in transfer is in need of rectification. The theory of the moral assessment of the past that I defend here is only committed to the claim that unjust acquisitions and transfers constituting cases of political violence are in need of rectification. As I said, the reason for this is the very fact that such cases flagrantly undermine the equal moral status of their targeted victims. Contrarily, not all unjust property acquisitions and transfers that Nozick's entitlement theory is set to rectify instantiate this quality. Another way of stating this caveat is the following: I defend that *at least* cases of past political violence are in need of rectification.

In this way my proposal is compatible with the claim that less stringent rectificatory measures might be taken (if any at all) regarding other types of historical injustice, but also with a stronger theory of rectification for other historical wrongs defended on a different basis.

A second caveat refers to my claim that adding a specific principle of historical rectification could fill justice-as-fairness's theoretical gap regarding the normative significance of historical injustice. While indeed a defence of the stated principle is offered throughout the thesis, a full development of it (and of its more general place within the framework of Rawls's theory) is not provided. As important as this task is within Rawlsian theory, this thesis is mainly concerned with showing how the assumption that historical injustice is outside the scope of Rawls's theory is unwarranted. If this thesis succeeds in showing that, it is up to further Rawlsian work either to fully develop the stated principle, and how it is to be placed within the framework of Rawls's theory more generally or, alternatively, to explain how Rawls's theory does address the normative significance of historical injustice without such a principle. I admit, then, that other Rawlsian strategies for securing historical rectification can be developed.

Finally, a third caveat is that, although there will be several remarks on the implications of the principle I am proposing, it is not the intention of this thesis to offer a particular theory of rectification. I hold that all actions of historical rectification can be subsumed to particular *retribution*, *compensation* or *recognition*-driven measures. Which specific measures of

rectification are to be deployed depends on the contextual elements of the case in question. It is not uncommon to find in the literature of transitional justice the idea that the three different measures just mentioned tend to clash with one other. One thing that I defend is that such putative tension is often overstated and misplaced. Relevantly, I hold that the importance of the third one cannot be underestimated in relation to the first two. For in *doing justice* to victims, recognition is as important as retribution and compensation. I also argue that rectifying cases of recent political violence requires of transitional frameworks including elements of the three types of rectificatory measures. When only one type of these measures is implemented most of the time rectification fails. However, as I have said, it is not the purpose of my thesis to develop a particular theory of rectification. Such theories must necessarily be sensitive to the particular contexts for which they are framed, and that is certainly out of the scope of what a general philosophical account of historical rectification within liberal societies is meant to provide.

Furthermore, all strategies that deal with specific cases of historical injustice are beyond the scope of what an ideal liberal theory, such as Rawls's, is meant to provide. Rather, what I intend to show is that Rawls's ideal theory, and especially his description of a well-ordered society, might be incomplete until a new principle concerned with historical injustice is added to it. This does not require showing *how* we can arrive from our current societies (in which most historical injustices remain un-rectified) to a

well-ordered liberal society (in which no relevant historical injustice remains un-rectified). Asking that of this thesis is tantamount to asking of Rawls's theory that it offer a particular strategy as to *how*, say, the basic institutions of the Mexican State can be effectively regulated by the principles of justice-as-fairness. You might find this task more important than merely discussing how an ideal well-ordered society should look. However, I do believe that a first step in guiding our real-world proceedings requires achieving normative clarity on what we want to accomplish. I think this is the spirit of Rawls's theoretical exercise to begin with. The relevance of my thesis in this regard is to show how current liberal theory, due to fashionable confusions, has failed to conceptualise correctly the importance of the public dimension of citizens' and liberal societies' pasts.

Part I

HISTORICAL RECTIFICATION AND LIBERAL THEORY

1

The Normative Significance of Historical Injustice

Past political violence has a tremendous impact on persons and liberal societies. In this chapter, I will present the empirical and normative evidence that confer relevance to its rectification. Attending both to what psychology and moral philosophy have to say, in § 1.1 I will present evidence granting the normative significance of historical rectification at a personal level. In particular, I claim that a non-rectified historical wrong may drastically hinder victims' capacities to achieve a healthy psychological life and self-respect. Attending both to what the social sciences and political philosophy have to say, in § 1.2 and § 1.3 I will present evidence granting the normative significance of historical rectification at a social level. In particular, I will claim that the rectification of recent as well as distant past wrongs is needed in order to restore citizens' moral relations and to safeguard the liberal integrity of social institutions.

1.1 Victims' Psychology and Self-Respect

Political violence is an institutional way of undermining the equal moral worth of persons. A person is selected as the target of this form of

wrongdoing because of her ethnicity, religion, culture, or political beliefs and actions. She can also be targeted for more than one of these characteristics at once (see Jones 2004: 2-10 and Verdeja 2006: 123). Through the lens of the aggressor, the victim is not worthy of equal moral consideration —if worthy of any moral consideration at all. The infliction of harm against her is an expression of this.

This form of wrongdoing is always directed, promoted or facilitated institutionally. Just as human rights cannot be violated by a single individual (see Pogge 2000: 47), a person committing an isolated hate crime is not thereby bringing about political violence. However, if the State fails to sanction several instances of similar hate crimes, this further failure does constitute political violence. In this case, while the State is not targeting wrongdoing against specific citizens directly, it is nevertheless allowing those citizens to be targeted without any consequences for aggressors.

Since political violence is an institutional way of undermining the equal moral worth of persons, it disrupts the most fundamental moral premises behind a liberal democracy. Regardless of the different theoretical approaches available to understand and justify such a political system, all of them agree that liberal-democratic institutions must guarantee that citizens are treated as persons with equal moral worth. John Rawls's theory of justice, for instance, begins by affirming this claim without further defence (see *PL*: 18-19, *JF*: 5). According to Rawls, a liberal democracy should be understood as a system of social cooperation between free and *equal* persons. Respect for

citizens' moral equality is thus assumed as one of the intrinsic features of a system of cooperation in order for it to count as liberal and democratic. In this thesis I will follow Rawls in this regard. No argument will be provided in support of this basic liberal-egalitarian premise.

In this chapter, I will try to show how non-rectified cases of past political violence (or, as I will refer to them, historical injustices) also disrupt the fundamental moral premises behind a liberal democracy. In this section, I will start by describing the importance of historical rectification on victims' individual lives. In the following sections, I will focus on the importance of historical rectification for liberal societies as a whole.

Not surprisingly, being a victim of political violence has severe effects on persons' capacity to achieve a healthy psychological adjustment as well as to develop successfully a plan of life. This is vastly documented by psychological studies. Studies also document the inverse effect: the beneficial impact of public rectification of historical injustice on victims' lives. During the 1990s, many psychologist and mental health workers thought that the positive effect of historical rectification was mainly due to the powerful therapeutic impact of testimony (see Agger & Jenssen 1996 —whose fieldwork took place during the Chilean transition to democracy in 1989-1991— and Wine 2006 —whose fieldwork took place in the aftermath of political violence in the Balkans as a result of the collapse of Yugoslavia).

However, further field studies have revealed that testimony is but one of many elements that help in the process of psychological recovery from

‘political trauma’. Psychologist Brandon Hamber 2006: 564, 2010: 97 —by focusing on post-*apartheid* cases in South Africa— stresses the positive effect that programs, objects, and actions of reparation in general (what he calls *reparations* in plural) have on citizens’ capacity for achieving a healthy psychological adjustment (what he calls *reparation* in singular) in the aftermath of political violence. These *reparations* go beyond the use of testimony:

The integral importance of reparations, remorse, restitutions, truth and acknowledgement to victims [...] I have found that participation by victims and survivors in processes aimed at achieving such elusive goals as truth and justice is an important component of healing —many survivors want to feel they are taking some action, even if they know it will not deliver complete justice or absolute truth. This gives survivors some control over their environment, something which political trauma normally overrides. (Hamber 2010: 194).

Each of these acts of rectification serves the purposes of *retribution*, *compensation*, or *recognition* for the victim of a past wrong. Bringing aggressors to justice may serve the purpose of retribution. Restitutions may serve the purpose of compensation. Testimony, remorse, and truth-seeking may serve the purpose of recognition. These three types of measures are now well regarded as highly effective on victims’ mental health. Socio-psychologists Brinton Lykes and Marcie Mersky state:

There is a general sense among mental health workers, psychosocial researchers and practitioners that all of these forms, including those that focus on material well-being, restoration of legal rights and property, judicial associations, truth-seeking processes, apologies, or institutional reform, can have important effects on psychosocial conditions at the individual and national or collective levels. (Lykes & Mersky 2006: 590-591).

It is essential to stress the public nature of all these strategies of historical rectification. Hamber 2006: 567 even states: ‘from an individual perspective,

reparations for human rights violations are trying to repair the irreparable'. But Hamber is not the first one to underscore the importance of active participation of society for the process of psychological healing. Many other studies now suggest that these public strategies to rectify historical injustice tend to have a good effect on victims' healing process *precisely* because political trauma itself has a social dimension, a dimension that is usually overlooked by the typical post-traumatic stress disorder (PTSD) clinical approach. This latter clinical approach focuses on the individual, diagnosing the pathology and then proposing therapy often based on the use of drugs and cognitive-behavioural interventions. By overlooking the social nature of political trauma, this type of therapeutic intervention fails to explain the efficacy of socially oriented therapeutic models.¹ Ignacio Martín-Baró — whose fieldwork took place in El Salvador during its entire civil war period (1980-1992)— summarises the roots of the inadequacy of PTSD therapeutic intervention:

The problem is rooted in the limited conception of human beings [of the PTSD clinical approach] [...] that denies their existence as historical beings whose life is developed and fulfilled in a complex web of social relations [...] To put it more plainly, mental health is a dimension of the relations between persons and groups more than an individual state. (Martin-Baró 1994: 109-11).

Martin-Baró's claim on the profoundly public character of mental health elucidates the public nature of political trauma. We can find similar

¹ Judith Herman herself may be an exception to this. In the work that inaugurated the new wave in psychological studies on political trauma, post-traumatic stress disorder, and recovery, she already notes how important it is for victims to actively engage in the social recognition of the harm done to them. Herman identifies three basic stages of recovery: establishing safety (Herman 1992: 155), reconstructing the trauma story (Herman 1992: 175), and restoring the connection between survivors and the community (Herman 1992: 196). In cases of political violence, all these stages require not only the active participation of

conclusions with regard to the importance of rectifying historical injustice as a way of pursuing communally oriented goals in the process of psychological healing in Backer et al 1995, Angger & Jensen 1996, Minow 2002, and Lykes & Mersky 2006. I will come back to the public nature of political trauma in § 1.2. For now, we only need to note the general tendency in psychological studies to regard methods of public historical rectification as beneficial for victims' mental health. In particular, from all these psychological studies, we can extract the following basic thesis: both the *social acknowledgement* of the terrible wrongs infringed upon victims as well as victims' *mere participation* in public processes of retribution, compensation, and recognition of such wrongs tend to be strongly beneficial to them in the struggle for achieving a healthy psychological adjustment.

By considering the importance of the rectification of historical injustice at this descriptive level, there are sufficient reasons for taking some public provisions to come to terms with an unjust past as one of the social bases of citizens' self-respect that the State must supply. As is well known, Rawls (*TJ*: 54, 386) refers to the social bases of self-respect as 'perhaps the most important primary good'. However, various authors have pointed out that Rawls seems to be ambiguous on whether self-respect is a psychological or a moral attitude (see Sach 1981, Massey 1983, Thomas 1983, Eyal 2005, *PL*: 404n), and on whether his notion is closer to an Aristotelian-oriented view related to the shame that a person may experience by failing to acquire excellence (as Rawls *TJ*: 386-391 himself notes) rather than to a Kantian-

oriented view related to persons' intrinsic worth as human beings (see Darwall 1977). This is not the place to consider at length such a notion and the debate it has generated. However, I would like to mention that an effective Rawlsian notion of self-respect (i.e., one that links the capital importance of self-respect to the capacity to pursue a plan of life) must include four elements.

The model I propose requires the fulfilment of four conditions in order to display self-respect. The first one, call it the *psychological condition*, states that the person must have a healthy self-esteem; that is, a positive evaluation of herself (see Constant & Foxx 2003). The second one, call it the *Kantian condition*, states that the person must value her status of equal worth with regard to the rest of the agents (see Hill 1973, Darwall 1977, and Meyers 1989). The third one, call it the *Rawlsian condition*, states that the person must try to realise the system of ends that she adopts (see *TJ*: 386, Hill 1985, and Raz 1986: 426). Finally, the fourth one, call it the *Aristotelian condition*, states that the person must demand that both her status as an agent with equal worth, as well as her own plan of life, are not devaluated by others (see Boxill 1976, Thomas 1983, Moody-Adams 1992, and *JF*: 85).

These conditions relate, respectively, to self-esteem, moral status, personal aspirations, and moral character. Because of this, none of them ensure that a person shows respect for herself when taken separately. This is manifest when we analyse the relation of self-respect with a person's system of ends.

The *psychological condition* states that the person must ascribe value to the plan of life that she adopts. For only when a person finds value in such a plan can she have a positive evaluation of herself. Conversely, a person who thinks that the ends she pursues are trivial might feel that her life as a whole has no real importance. This feeling can be exacerbated by the repeated experience of shame and failure in achieving previous and now abandoned key ends. Hill's famous *self-deprecator case* exemplifies this (see Hill 1973: 88 and Dillon 1992: 126-127).

The fulfilment of the *psychological condition* would prevent this from happening. However, since an agent can endorse a plan of life highly damaging for herself, having a positive evaluation of such a plan is not sufficient for displaying self-respect. This is the case of ways of living in which a person positively endorses systematic humiliation (for instance, a beggar who walks around the street targeting persons and then strategically focuses on her multiple misfortunes up to the point of sincere crying every time she asks for money). Such a life style is based on the systematic undermining of self-respect, regardless of whether agents can positively endorse it in some circumstances.

Here the *Kantian condition* complements the *psychological condition*, since it prescribes that the person must refrain from following plans of life that diminish her equal moral worth. However, since the *Kantian condition* only highlights a negative aspect in the election of plans (by stating which plans should *not* be endorsed), there is still the possibility that even when the

person positively endorses a plan of life which respects her equal worth, she does nothing to achieve it. Hill's famous *deferential wife case* (see Hill 1973: 88) is an illustration of this. In this case, despite the tremendous importance that the person ascribes to her ends, she simply feels incapable of achieving them. Because of this putative incapacity, the deferential wife ends up believing that she should sacrifice all of her ends in support of the fulfilment of her husband's ends.

Here the *Rawlsian condition* complements the previous two, since it highlights a positive aspect in the election of plans: the person must try to achieve the plans she adopts. However, as in the former cases, even when the *Rawlsian condition* is taken in conjunction with the previous two, there is still the possibility of a person displaying a deficit in self-respect. For, even when she adopts a plan of life to which she ascribes value that respects her equal moral worth, and when she tries to achieve such a plan, it may be that she cannot effectively do so because of the interference of external factors such as the direct action of other agents or through the mediation of the State. Here the *Aristotelian condition* complements the previous ones, for it states that the person must demand of others respect for her plan of life.

Note three things on the account of self-respect just proposed. First, that the account is descriptive: it tells when a person shows respect for herself and not that a person *ought* to respect herself or that she is to be *blamed* for not displaying self-respect (see Dillon 1992: 125). Second, while I do believe that each of the four conditions is necessary for displaying self-

respect, I do not claim that they are jointly sufficient. Thus, the account leaves it open as to whether a person who disrespects others shows thereby a lack of self-respect (see Buss 1999: 540). Third, the account also leaves it open as to whether only a person that respects herself *can* respect others (see Verdeja 2006: 118).

One thing that is quite clear in Rawls's notion of self-respect is its relation to the effective pursuit of a plan of life —see (*TJ*: 386). By complementing such a notion with the conditions that I have described, we can see why the social bases of self-respect are the most important primary good. For instance, when the State offers provisions to come to terms with a past shaped by political violence, it is thereby helping victims to achieve self-respect in each of these four levels: self-esteem, moral status, personal aspirations, and moral character.

However, as Rawls (*JF*: 60n) himself does, it is important to note that the primary good in question is *the social bases* for achieving self-respect, rather than self-respect itself. The rationale for this is twofold. First, as it is well known, the metric of primary goods must provide a public basis of interpersonal comparisons —see (*PL*: 178-187)— and only the objective *means* to achieve self-respect can have this characteristic. Second, although the State can provide the necessary means for citizens to being able to achieve self-respect, it cannot guarantee that every single citizen will respect herself. This is particularly clear with the social basis I am proposing: even if the State was to engage in the rectification of past wrongs, that would not

guarantee a full recovery of their victims (see Minow 2002 and Hamber 2006).

The fact that Rawls does not regard rectification of historical injustice as a key element for achieving self-respect highlights a limitation in his conception of the person: that such a conception is mainly forward-looking in character. Very few Rawlsian passages relate to the importance of the past in persons' life. Consider, for instance, what Rawls claims in the context of explaining why a well-ordered society can be understood as a 'social union of social unions' (*TJ*: 462) —that is, as an institutional arrangement in which the good related to the sociability of humankind is realised:

To say that man is a historical human being is to say that the realizations of the powers of human individuals living at any given time takes the cooperation of many generations (or even societies) over a long period of time. It also implies that this cooperation is guided at any moment by an understanding of what has been done in the past as it is interpreted by social tradition. (*TJ*: 460)

While this passage explicitly emphasises the importance of the past for a person's life, it does so more in the way of showing how every human endeavour is historically placed and thus dependant on the social tradition to which each person belongs. It is rare to find more specific claims regarding the importance that the particular and individual past of a given person holds for herself. Concerned with persons' capacity to effectively achieve a plan of life, for the most part, Rawls's conception of the person has a strong forward-looking orientation.

However, the evidence against the plausibility of adopting such an exclusive orientation is abundant. The importance of the individual past in a

person's life is also well documented beyond the literature on recovery from political trauma. For instance, numerous psychological studies —ranging from empirical psychology (see Neimeyer & Metzler 1994) to cognitive psychology (see Neisser 1994 and Conway & Pleydell-Pearce 2000) to neuroscience (see McGaught 2003, Fivush & Nelson 2004, Svoboda, E. *et al.* 2006)— emphasise the mechanisms of autobiographical memory (i.e. recollection of personal past events) in a person's self-understanding.

Neuroscientist James McGaught says in the opening passages of his most recent book: 'All of our knowledge of the world, and our skills in living in it, are based on memories of our experiences. So, too, are all of our plans and dreams. Life without memory is difficult to imagine [...] A life without memory would be no life at all' (McGaught 2003: ix). McGaught's remarks here may seem even platitudinous. This stresses how we tend to regard as completely inaccurate a forward-looking conception of the person; for, as he says, 'memory is the "glue" of our existence' (McGaught 2003: 2). In this sense, some experimental work on autobiographical memory shows the intricate relationship between personal identity (characterised as self-appreciation) and memory, illustrating that, as Neimeyer & Metzler 1994: 105 hold, 'our identities and memories are two sides of the same coin'.

Conway & Pleydell-Pearce state:

Autobiographical memory is of fundamental significance for the self, for emotions, and for the experience of personhood, that is, the experience of enduring as an individual, in a culture over time. (Conway & Pleydell-Pearce 2000: 261)

The processes of self-identification and autobiographical memory are also developmental. Some developmental psychologists even think that the links between self-identification and autobiographical memory go as far back as the preschool years (see Fivush & Nelson 2004: 486-487). From all this psychological work we can extract the following basic thesis: without bringing recollections of our past into a somewhat particular and coherent form, we would simply not see ourselves as we do. In other words, our self-understanding at any given time has a strong backward-looking component.

The work of some philosophers also adds to the case against Rawls's forward-looking conception of the person. Christine Korsgaard (see Korsgaard 1996a: 363-398), for instance, points at two presuppositions of effective agency: in order to elect any given course of action, the agent must think of herself both as a *synchronic unity* (at a given time) and a *diachronic unity* (over time) of concern. Regarding the former, the agent must suppose that at the time of an election she is one single unity who can bring about choices and actions that may be beneficial or prejudicial to such a unity. Regarding the latter, Korsgaard says:

[M]ost of the things we do that matter to us take up time. Some of the things we do are intelligible only in the context of projects that extend over long periods. This is especially true of the pursuit of our ultimate ends. In choosing our careers, and pursuing our friendships and family lives, we both presuppose and construct a continuity of identity and of agency. (Korsgaard 1996a: 371)²

² Many other authors have defended the diachronic character of agency. Harry Frankfurt 1999: 139 holds: 'there is, I believe, a quite primitive human need to establish and to maintain volitional unity. Any threat to that unity —that is, any treat to the cohesion of the self— tends to alarm a person, and to mobilize him for an attempt to "self-preservation"'. Charles Taylor 1989: 52 states that unity over time 'is an inescapable structural requirement of agency'. John Campbell 1994: 190 holds that temporal 'identity is central to what we care

These remarks about the diachronic unity of agency take place in a discussion with Derek Parfit on the implications that the nature of personal identity may have for rational and moral choices —see Parfit 1984: 307-345. This may explain why Korsgaard only points at the *forward-looking* component of such a diachronic unity: in order to make choices and, at the limit, to elect a Rawlsian plan of life, an agent must suppose that she will be the same person as the one who in the future will enjoy such a plan. However, there is also a clear *backward-looking* component of the diachronic unity of agency: for *the agent's past life is the basis of all of her elections*. In order to elect a plan of life, the agent must take into account the aspiration she has previously formed for herself, and more generally, her past experiences, decisions, attachments, commitments, etc. There is no other way of electing one plan of life over others. This is the case even when the agent wants to undergo a radical personal transformation by committing herself to a completely renewed plan of life. We think that such a new plan of life is a *radical transformation* precisely because of a comparison with the previous personal path.

The backward-looking components of self-understanding —on the one hand— and of agency —on the other— highlight an important deficit in Rawls's forward-looking conception of the person. Moreover, as we have seen, the relation of the backward-looking component of agency to self-

about: one thing I care about is what I have made of my life'. Alasdair MacIntyre 1981: 203 states that, 'the unity of an individual life is the unity of a narrative embodied in a single life. To ask "What is good for me" is to ask how best I live out that unity and bring it to completion.'

respect is patent. For a past marked by extreme abuse and unfairness may have severe effects on the capacity to successfully achieve a plan of life. As we saw, such a past may strongly diminish agents' present self-esteem, moral status, personal aspirations, and moral character.³

Notice, however, that I am not claiming that there is no single victim who may achieve a healthy psychological adjustment and live a meaningful and cooperative life in society without coming to terms with the past. This is already documented in Hamber 2006: 568: 'Degrees of dealing with the consequences of extreme political violence and trauma are possible. Many victims are indeed survivors and highly resilient'. An incredible example in this regard is that of the current president of Uruguay, José Mujica, who during the 1970s and 1980s spent a total of fourteen years in prison, eleven of them in constant torture and strict seclusion as one of the chief guerrilla 'hostages' of the military junta that governed Uruguay from 1973 to 1985.

³ I have defended that both psychology and moral philosophy grant the importance of the past in a person's self-understanding and agency. Galen Strawson attacks a characterisation of both claims (see Strawson 2004). Although I cannot develop this in depth here, I find Strawson's attacks implausible. There seem to be two main motivations behind Strawson's critiques. First, that it is not a metaphysical indisputable truth that persons are diachronic units of concern (see Strawson 2004: 428). Second, that there are good ways to live in which not much weight is assigned to the past (see Strawson 2004: 429). My remarks on the diachronic unity of self-understanding and agency are compatible with these two claims. While I will take issue with the former in § 2.3, I will concede the latter by mentioning cases of highly resilient people that in fact seem to enjoy living a forward-looking life. In this sense, Strawson does not consider that psychological facts tend to work as human regularities or tendencies, not as necessary phenomena experienced by every single human being (see Baldwin 2008). In fact, it is somewhat surprising that Strawson does not mention one single psychological study on autobiographical memory, despite the fact that one of his two main targets is a thesis held by researchers in that area of psychology. The two putative motivations of Strawson's critiques are explicitly addressed by Neimeyer & Metzler 1994. These authors note correlation patterns between different personal characters and different forms of personal recollection: just as Strawson is so eager to prove, some people whose identity is less steady at a given time tend to be much more flexible regarding the understanding and significance of certain facts about their past —see Neimeyer & Metzler 1994: 128.

Despite this past shaped by extreme brutality, Mujica managed to become president of his country about twenty years later and has repeatedly stated that it is not his personal goal to open processes of retribution, compensation, and recognition for such and similar past injustices.⁴

I will come back to the case of Uruguay and Mujica in § 3.1. For now, my claim is simply that even when *some* victims can achieve a forward-looking life like Mujica's, a theory of liberal justice cannot ask (due to the evidence reported here) of *all* victims to live their lives in a strictly forward-looking way. In this sense, my argument is very similar to one analogical argument offered by Will Kymlicka for taking as a primary good the access to the liberally compatible ways of life available in one's culture:

Liberals rightly assume that the desire for nonsubsistence resources is so normal —and the cost of forgoing them so high for most people's way of life—that people cannot reasonably be *expected* to go without such resources, even if a few people voluntarily choose to do so [...] Similarly, I believe that, in developing a theory of justice, we should treat access to one's culture as something that people can be expected to want, whatever their more particular conception of the good. Leaving one's culture, while possible, is best seen as renouncing something to which one is reasonably entitled. (Kymlicka 1995: 86. *Italics in the original*)

Just as Kymlicka claims that the cost of adapting to a different context of choice cannot be imposed over citizens of minority cultures, so I think that the cost of trying to live a forward-looking life cannot be imposed on citizens that were victim of past injustice.⁵

⁴ Mujica recently said in an interview: 'As a person, I do not like to live looking back, because life is always to come and every day dawns. But that is the way I am. I cannot impose that on my co-citizens'. See "Entrevista a José Mujica", *El País* 17/04/2011. (Translation mine)

⁵ The pertinence of this analogical claim does not depend on whether Kymlicka is right in thinking that the access to the liberally compatible ways of life of one's culture should count as a primary good. I bring this passage to attention because of the *structural* similarity of Kymlicka's claim and mine. For the purposes of this thesis it is irrelevant whether Kymlicka's

1.2 Restoring Mutual Respect and *Recent-Past Rectification*

Liberal societies observe a social regularity: they tend to eventually engage in processes of *recent-past rectification*. Almost all contemporary transitional liberal societies embark on processes aiming at coming to terms with their recent unjust past.

José Zalaquett —ex-president of the Inter-American Commission of Human Rights, of Amnesty International, and a member of the Chilean Truth and Reconciliation National Commission in 1990—, has detailed over the years how, ever since the reestablishment of civil government in Argentina in 1983, the world has seen a wave of *truth* and *reconciliation* commissions dealing with past injustices on all continents. Truth commissions or similar institutional bodies have been formed in Bolivia, Brazil, Chile, East Timor, Ecuador, El Salvador, Fiji, Ghana, Greece, Guatemala, Haiti, Liberia, Mexico, Morocco, Panama, Peru, The Philippines, Sierra Leone, Solomon Islands, South Africa, South Korea, Sri Lanka, and Uganda (see Zalaquett 1998, 1999, and Grandin & Miller 2007).

Of course, it would be a mistake to think that this regularity in liberal societies has so far ensured retribution, compensation, and recognition for victims. Rather, the work of truth commissions and similar bodies around the world appears as an unfinished business in the great majority of cases. To begin with, most of these commissions have strict limits in their capacities. They tend to lack the judicial capacity to prosecute and convict past

claim on access to culture as a primary good is incorrect or incorrect. I will say more about the stated structural similarity at the very end of § 2.1.

aggressors. Moreover, when they do have this capacity, or when independent judicial procedures are undertaken, the results tend to be highly questionable (see Elster 2004: 47-66).

Sometimes it is argued that these poor results are in part due to the fact that *retributive* and *recognition*-driven measures clash with each. It is not uncommon to find in the literature on transitional justice the claim that the aims of retributive justice and truth grow apart (see Van Zyl 1999, Elster 2004: 116, Lutz 2007: 326, Smyth 2007: 6-21, Roper and Barria 2009).

Scholar Marie Breen Smyth states:

Whilst the provisions of amnesty may be a valuable incentive to perpetrators to participate in a truth recovery process, the provision of amnesty precludes the achievement of justice for victims [...] Victims in general are seemingly faced with a choice between more information or 'truth' on the one hand and justice on the other. (Smyth 2007: 13)

The putative conflict between justice and truth is often overstated and misplaced. We should not think that the focus on truth of rectificatory institutional bodies is solely based on aggressors' political hopes for a lack of real accountability. For the importance of truth-seeking relates to the acknowledgment of past wrongs as such and thus to the recognition that terrible things were perpetrated directly on victims. We should thus note that in past injustice rectification, that is, in *doing justice* to victims, recognition is just as important as retribution.⁶

Likewise, another frequent way of understanding the work of recent-past rectification programs is through distinguishing between *material* and

⁶ The conflict between *retribution* and *recognition*-driven measures will be the focus of § 3.3.

symbolic forms of reparation (see Elster 2004: 329, de Grieff 2006: 467-471, Lira 2006: 65-66, Verdeja 2006: 123). Whereas compensatory schemes providing direct economic remuneration to victims of past wrongdoing are usually listed as the most clear forms of material reparation, truth telling mechanisms and public apologies are listed as some of the most common forms of symbolic reparation. Yet two reasons show how this is another mistake in the literature on transitional justice.

First, as Margaret Walker 2010: 530 has defended, *all* acts of rectification are symbolic. When one has been tortured and imprisoned because of one's political beliefs, as was the case for thousands of citizens in Argentina during the military rule of 1976-1983 (see Guembe 2006: 25-47), economic compensation is nothing more than a symbolic and extremely limited measure of institutional acknowledgment of wrong.

Second, the stated division suggests that economic compensation constitutes the only tangible way of rectifying past political violence, as if the public acknowledgment of wrongdoing were an intangible, immaterial, fake, way of rectification. However, this characterisation conflicts with the way in which several victims respond to economic compensation for past political violence. Victims sometimes perceive that this rectificatory measure is a form of 'blood money'; that is, a way of silencing their claims for justice in exchange for economic advantage. In fact, this is the case of many Chilean citizens who, after losing their relatives in illegal detentions and forced disappearances under Chile's brutal military rule of 1973-1990, qualify as

being the beneficiaries of a life-long monthly pension. While some of these citizens refuse to accept the pensions (see Elster 2004: 166n, Lira 2006: 93), many of them now forcefully complaining about the government's proposal to stop referring to Augusto Pinochet's regime as a 'Dictatorship' in primary-school textbooks.⁷ So, again, we should note that in past injustice rectification, that is, in *doing justice* to victims, recognition is just as important as compensation.

But regardless of the stated problems in most processes of recent-past rectification—which include questionable or insufficient *retribution*, *compensation* or *recognition*-driven measures—, it is a well-documented social regularity that most liberal societies eventually undertake such processes.

In this sense, recent-past rectification, when successful, serves very important normative goals. Aside from the more personal benefits reported in § 1.1, processes aimed at coming to terms with a recent unjust past also accomplish important *public* goals. This is linked to the public nature of political violence. Lykes and Mersky 2006: 591-592 hold that: 'repair from political violence must be distinguished from repair from the psychological distresses caused by severe natural disasters or even as the result of an individual criminal action. The former has a singular moral dimension'. The singular moral dimension of repair from political violence has to do with two things: contrary to natural disasters and individual criminal action, political

⁷ See 'Pinochet "dictatorship" textbook row erupts in Chile', <http://www.bbc.co.uk/news/world-latin-america-16420413>.

violence is always targeted. As we have seen, a person is selected as the target of this form of wrongdoing because of her ethnicity, religion, culture, or political beliefs and actions. Also, and very important for the social dimension of rectification, contrary to individual criminal action, political violence breaks the moral standing of the victim not only with regard to her aggressor, but also with regard to the rest of her community.

Considering these two elements, the normative significance rectifying past political violence can be framed in terms of the respect that society owes to its victims. Showing respect for a person requires acknowledging her equal moral worth; that is, recognising that her life and system of ends is due equal consideration because of the simple fact that she is a person. Since, as we have seen, political violence is an institutional way of undermining victims' equal moral worth, full respect for such victims can only be restored by the adoption of institutional measures directed to rectify this past failure (see Kutz 2004: 284).

In this sense, all rectificatory measures are primarily a way of acknowledging that institutions have failed to show respect for some of their citizens in the past. The first thing needed to achieve this goal is to assert publicly and with no reservation that what victims have suffered was wrong and should not have occurred. It could be argued that this is already known in a society that has suffered from political violence in the past, and thus that no new knowledge would be gained by establishing public mechanisms, such as truth commissions, for stating it. Yet we should notice the difference

between knowledge and acknowledgement proposed by Thomas Nagel: 'acknowledgment [...] is what happens and can only happen to knowledge when it becomes officially sanctioned, when it is made part of the public cognitive scene' (quoted in Weschler 1989). Making the knowledge of past injustice part of the public cognitive scene, as I have been arguing, is a crucial step in restoring respect for victims.

Now, when a society radically fails to respect some of its members, public trust in its institutions is shattered (see de Grief 2006b: 460). Victims might not deposit trust in a social arrangement that does not condemn its own failure to safeguard their equal moral worth. Furthermore, non-victims can develop this attitude as well. So long as no explicit disapproval is expressed regarding the previous deprivations conducted or facilitated by the State, citizens in general might believe that their institutions do not hold a real commitment to the fundamental moral premises of liberal justice.

Indeed, this is the case of many Latin American countries in the present (see Verdeja 2006: 134). In these countries, the incredibly low trust deposited in social institutions is not only due to the fact that numerous irregularities are committed by present political actors, it is also due to the fact that *it has always been this way*. Some perpetrators of gross past crimes walk with full impunity and continue to enjoy life-long pensions for their term in office—as it is the case, for instance, of the Mexican ex-president Luis Echeverría (1970-1976), who has been repeatedly identified as responsible for the direct persecution and assassination of Marxist militants, organised

indigenous groups, and students during his term in office (see Acosta & Elennin 2006: 108). Against this scenario, historical rectification is a way of encouraging trust in the liberal commitment of current institutions. For, by acknowledging the wrong done to victims as a *wrong*, as something that should not have occurred, institutions make it clear they *reject* a non-liberal form of social interaction; that is, that they reject a social context in which persons are not treated as equals due to such characteristics as political beliefs or cultural, ethnic or religious belonging.

Because of the aforementioned reasons, a historical injustice not only affects victims' psychology and *self*-respect (as we saw in § 1.1), but also the *mutual* respect that citizens owe to each other in line with the proper morality behind social interactions in a liberal society. In this sense, a historical injustice follows the description given by Thomas Scanlon regarding the two different 'evils' that an unjust inequality might create.

Scanlon holds:

"It is an evil to be treated as inferior" [...] The experiential evil involved here can be characterized in several different ways [...] Let me distinguish two broad categories. The first, more "individualistic", characterization emphasizes what might be called damage to individuals' sense of self-worth: such things as feelings of inferiority and even shame resulting from the belief that one's life, abilities and accomplishments lack worth or are greatly inferior to those of others. The second category emphasizes damage to the bonds between people: what might be called the loss of fraternity resulting from great differences in material circumstances, accomplishments and the social importance according to them. (Scanlon 2003: 212)

Scanlon's description of the two types of evils that an unjust inequality produces is quite significant. While the first one relates to the damage

inflicted on the self-respect of the person affected by the inequality, the second one has to do with to the relation between that person and the rest.

The fact that unjust inequalities, just as historical injustices, create these two distinct normative affections reflects that *self*-respect and *mutual* respect (or fraternity, as Scanlon calls it) must be distinguished. At times Rawls seems to collapse the second one into the first one (see *TJ*: 155-156, 297, 478), as if all deficits in mutual respect would bring about a deficit in self-respect.

Colin Bird 2008: 17 refers to this latter claim as ‘the dependency thesis’ —i.e., the claim that all deficits in mutual respect bring about a deficit in self-respect. However, the dependency thesis should be rejected. For once we follow Rawls on the main idea that a healthy self-respect is connected to the successful pursuing of a plan of life (see *TJ*: 386), the independence between the two forms of ‘respect’ is evident: one might successfully achieve a plan of life *even* while experiencing some important deficits in mutual respect. This is the case of many victims of historical injustice. Despite the fact that the State fails to show respect for them (by failing to offer institutional provisions for historical rectification), they nevertheless manage to successfully develop a meaningful plan of life with which they fully identify.

Conversely, a person might fail to display self-respect *even* when experiencing full respect by society. This is part of the rationale for taking the *social bases* for self-respect as a primary good rather than self-respect itself.

When these social bases are secured, the State meets its obligation of showing respect for citizens. However, as I have said, it might still be the case that a person experiences a deficit in self-respect even when these bases are provided. This can have many causes. In the case of victims of past political violence, the distress of having experienced gross forms of wrongdoing can simply be unbearable (see Hamber 2006: 567), even if rectificatory measures including full and proper psychological counselling are in place (as part of the social bases of self-respect that the State provides).

This shows that, *pace* Margalit 1994: 124, whether a person displays respect for herself or not is *ultimately* dependent on her own strength and capacities. It depends on whether she manages to fulfil the four conditions proposed in the previous section: having a positive evaluation of herself (psychological condition), endorsing a system of ends compatible with her equal moral worth (Kantian condition), trying to fulfil such ends (Rawlsian condition), and demanding that others not prevent her from achieving them (Aristotelian condition).

Conversely, whether a person is treated as an equal, as mutual respect commands, is not *ultimately* dependent on her. In the case of victims of past political violence, as we have seen, this ultimately depends on whether the institutions of the State provides means for historical rectification or not.

1.3 Liberal Integrity and *Critical Historical Narratives*

Liberal societies observe a further regularity: all of them embark on the creation of *historical narratives*. *Official* history is always told, taught, and promoted by the institutional means of the State in a specific way. Although there is a plurality of competing versions of a given society's history, some of them created for instance at academic institutions or political think tanks, *no liberal society leaves this job exclusively in private hands*. Inevitably, certain facts about the putative origin, distant, and recent past of society are highlighted and others omitted by the use of public institutional means.

One initial reason why this is so is that there are practical limits regarding the means for registering the past: no historical narrative can include all facts, since there are simply not sufficient historiographical elements to do so. Most parts of the past as such are lost.⁸ More importantly, the selectivity of historical narratives regarding the inclusion of facts is connected to the main reason why liberal societies construct such narratives in the first place: i.e., the fact that each society is both a *cultural* and a *political* continuum over time.

That liberal societies are *cultural* communities extended through time has been defended by authors sometimes identified as 'liberal nationalists'

⁸ Arthur Danto 1985 offers an alternative defence of this thesis by stating that all true descriptions of historical facts are also historical facts. As some of these descriptions of historical facts can be about events yet to come (what Danto calls *narrative sentences* such as —supposing for the sake of the argument that the following is true— 'Barack Obama will be re-elected in 2012 and will repeat his *de facto* amnesty policy towards massive financial fraudsters'), it is *impossible* at any given time to offer a complete account of historical facts. Following this rationale, Danto 1985: 17-18 states: 'Any account of the past is *essentially* incomplete' because 'a complete account of the past will suppose a complete account of the future'.

(such as Yael Tamir 1993, David Miller 1995 and Will Kymlicka 1989, 1995, 2007).⁹ These authors have emphasised how liberal societies form and promote a collective cultural identity through the State's institutional means (i.e., by the implementation of official languages, official dates and holidays, forms of civic education, and historical narratives). This is a claim with vast empirical support.

That liberal societies are *political* communities extended through time has been defended by Rawls himself. In fact, this commitment is expressed in what Rawls calls the 'most fundamental idea' of his conception of justice: i.e., that a liberal society must be understood as a 'fair system of social cooperation over time from one generation to the next' (*JF*: 5). Rawls repeatedly states this idea: 'Recall that a political society is always regarded as a schema of cooperation over time indefinitely' (*JF*: 162). In similar fashion, Janna Thompson defends that political communities have intergenerational span: 'a polity is by nature intergenerational. It has intergenerational responsibilities. It has institutions that are essentially intergenerational' (Thompson 2009: 12).

⁹ Several (usually overlooked) passages of Rawls's theory support the claim that liberal societies are cultural communities extended through time. For instance: 'For normally leaving one's country is a grave step. It involves leaving the society and culture in which we have been raised, the society and culture whose language we use in speech and thought to express ourselves, our aims, goals, and values; the society and culture whose history, customs, and conventions we depend on to find our place in the social world. In large part we affirm our society and culture, and have an intimate and inexpressible knowledge of it, even though much of it we may question, if not reject' (*PL*: 222). Consider also the passage quoted in § 1.1: 'To say that man is a historical human being is to say that the realizations of the powers of human individuals living at any given time takes the cooperation of many generations (or even societies) over a long period of time. It also implies that this cooperation is guided at any moment by an understanding of what has been done in the past as it is interpreted by social tradition' (*TJ*: 460).

These two features explain the creation of historical narratives by the use of institutional means: such narratives are the way in which societies make sense of their temporal span as cultural and political continuums. The importance of a historical narrative is thus directly connected with its links to the present: society's own sense of identity throughout time is dependent on it. Therefore, focusing on certain facts while omitting others is never left to chance.

I believe that this institutional exercise should meet a normative standard: the narrative promoted by the State must be, as I will call it, *critical*. A critical historical narrative —for short, CHN— is a particular understanding of the past that is sensitive to the relevant historical injustices that occurred within the temporal span that the narrative comprehends.

There is one fundamental reason to support the claim that a liberal society should promote a CHN. Remember, once again, that political violence is an institutional way of undermining the moral worth of persons. Since, in turn, a liberal democracy is strongly committed to the protection of persons' equal moral status, an institutional account of the past that *systematically celebrates* or *omits* the occurrence of past political violence is incompatible with a liberal democracy. Narratives so constructed, regardless of how common they are in the present day, shatter the liberal integrity of the societies adopting them (see Abizadeh 2004: 309). Contrarily, a CHN is a form of historical rectification that promotes such integrity. Just as official provisions for recent-past rectification, a CHN is a way for institutions to

state their condemnation of the non-liberal social context that allowed the perpetration of political violence in the past.

Considering this, it could be argued that part of the duty that — according to Rawls (*TJ*: 99)— citizens have to uphold, perpetuate, and defend just institutions, requires that citizens support institutions which promote a CHN (see Thompson 2006: 162-163). Since, as we have seen, liberal societies are institutionally continuous over time, citizens may not only be concerned with their institutions' performance in the present. For instance, they may find it disturbing that in the evolution of such structures up to the present, several departures from the principles that are to guide just institutions took place. Of course, they cannot set the course of their society's distant past in line with the precepts of justice. However, since their present institutions will promote a particular understanding of such a past by means of the creation of a historical narrative, they can nevertheless demand such a narrative be critical. This, it seems to me, is the best way of discharging their duty to uphold just institutions in this case.

Sometimes the existence of a similar duty is vindicated by the fact that citizens take *pride* on the achievements of their cultural group even as far back as its putative historical origins; they also tend to *resent* offences to national ancestors committed by other nations as their own.¹⁰ Considering this, Farid Abdel-Nour has recently defended that:

¹⁰ Michael Ignatieff offers a neat example in this regard: 'reporters in the Balkan wars often observed that when they were told atrocity stories they were occasionally uncertain

We can then state that to national belonging, the dominant form of collective political identification in the last two hundred years, is attached a concomitant responsibility. The agent who participates in this form of belonging incurs a responsibility for that part of herself that is caught up in the nation. This responsibility she incurs by means of her national identity alone [...] My conclusion is simple. Where there is national pride, there is national responsibility. The latter can only disappear in a world devoid of national identity. Until then, let the participants in national identity recognize that their fantasy opens the door to “the guilt of the fathers.” (Abdel-Nour 2003: 712-713)

Abdel-Nour holds that national identification gives rise to national responsibility —similar remarks have been defended by Kutz 2004: 279-285 and Butt 2009: 178-188. In contrast, Thompson 2006: 157 has criticised this approach by highlighting the conditionality of Abdel-Nour’s claim: even when a great amount of citizens tend to experience national pride, some simply do not. These citizens might not identify at all with the putative historical landmarks of the cultural group they belong to.

While I think Thompson is right in emphasising the conditionality of this ‘nationalist’ way of defending the need of a CHN, she fails to see that, regardless of whether citizens identify on a personal level with the history of their cultural group, the State will nevertheless promote a historical account of the past highlighting not only society’s *institutional* continuity through time but also strong forms of *cultural* continuity. This is, precisely, what liberal nationalists repeatedly underscore. Now, since the State is a figure for citizens’ institutional *representation*, when it promotes a particular understanding of the past that is blind to the relevant cases of past political violence, it does so *in the name* of its own citizens. The State thereby makes

whether these atrocities had occurred yesterday or in 1941, or 1841, or 1441’ (quoted in Minow 1998: 14).

all citizens collaborators in the process of normative irregularity affecting a society that lacks a CHN, regardless of whether they identify with such a narrative or not.¹¹ This is, thus, a further reason in favour of the idea that citizens should support institutions that promote a CHN: their own liberal integrity is compromised when the State adopts a non-critical historical narrative.

Sometimes it is suggested that historical narratives should not be informed by a criterion of truth because only in this way they can achieve the goal of enhancing solidarity, stability, or even civic education. Let me address these concerns in turn. Consider what Miller says:

Once we discover that national identities contain elements of myth, we should ask what part these myths play in building and sustaining nations. For it may not be rational to discard beliefs, even if they are, strictly speaking, false when they can be shown to contribute significantly to the support of valuable social relations. (Miller 1995: 35-36)

According to Miller, constructing a historical narrative without the constraint of a criterion of truth might help to enhance social unity and stability. He nevertheless makes it clear that this might not hinder the adoption of a CHN since, according to him, the construction of a historical narrative must be open to a 'process of debate and discussion to which everyone is potentially a contributor' (Miller 1995: 36). This ensures that historically oppressed groups could participate in such debate and thus demand that the injustices that their forebears suffered are included in the narrative.

¹¹ In this sense, Rawls (*PL*: 216) holds that 'in a democratic political power, which is always coercive power, is the power of the public, that is, of free and equal citizens as a collective body'. In a liberal society, the coercive use of the State's power is thus connected to each citizen. For an interesting analysis on the relationship between institutional injustice and citizens' individual responsibility see Pasternak 2011.

But even if we assume that something close to a CHN can still be produced without a criterion of truth, Miller's rejection of such a criterion is highly problematic. For Miller makes the adoption of a particular historical narrative a matter of political compromise between different social groups — that is, what Rawls (*PL*: 147) calls a *modus vivendi*. In contrast, the importance of adopting a CHN is not that such an understanding of the past would please every relevant cultural group in society but, rather, that some extremely important cases of past political violence *occurred* in the past and that this is need of acknowledgement. A criterion of truth is thus a necessary constraint of a CHN. Ultimately, a CHN is an institutional form of historical rectification, and rectification for an injustice that did not occur is nonsensical. So, even if it is certainly the case that a CHN will promote social stability in a highly plural and multicultural society, the adoption of such a narrative should be based on the importance of acknowledging past wrongdoing. This shows that a CHN promotes social stability *for the right reasons*.

As it is well known, the goal of social stability has a prominent role in Rawls's theory. The body of ideas compiled in his political liberalism was introduced into his original theory as a way of dealing, precisely, with what he saw as a problem with the stability of a society regulated by the principles of justice-as-fairness (see *PL*: xvi). The prominence of this goal is also stressed by Rawls's famous dictum: 'Let us agree that a political conception must be practicable, fall under the realm of the possible. This contrasts with a

moral conception that is not political: a moral conception may condemn the world and human nature as too corrupt to be moved by its precepts and ideas' (*JF*: 185). In light of this, Rawls defends that social stability for the right reason requires the moral acceptance of the principles of justice-as-fairness. In similar fashion, I hold that social stability for the right reasons requires the promotion of a CHN.

The importance of this latter claim is manifest by considering the fact that, as Kymlicka has consistently defended, no contemporary liberal society is a culturally homogeneous political community (see Kymlicka 1989: 206-220, 1995: 10-26). All liberal societies, aside from observing pluralism of comprehensive doctrines, observe strong forms of cultural pluralism: they contain national minorities (such as Scotland within the United Kingdom), indigenous communities (such as the Inuit in Canada), and immigrant populations (such as the vast Mexican community in the United States). A historical narrative publicised with shared institutional means that is unresponsive to the *facts* regarding these groups' past interactions might damage the social bonds between them. However, as I have said, we need to keep in mind that a CHN promotes such bonds because of the right reasons (i.e., acknowledgement of past mistreatment according to liberal standards).

Let me turn to the relation between historical narratives and civic education. Very few liberals would hold that civic education should not be provided by the State. According to Rawls, for instance, the State should promote the *political virtues* that must shape citizens' moral character. These

include 'the virtues of social cooperation such as civility and tolerance, of reasonableness and the sense of fairness' (*PL*: 194). Other scholars believe that such an education should be far more comprehensive than what Rawls's political liberalism is willing to concede. This is the case of William Galston 1994. Oddly, Galston thinks that this more comprehensive civic education demands, amongst many other things, the institutional promotion of a non-critical historical narrative:

Rigorous historical research will almost certainly vindicate complex "revisionist" accounts of key figures in American history. Civic education, however, requires a nobler, moralizing history: a pantheon of heroes who confer legitimacy on central institutions and are worthy of emulation. (Galston 1991: 244)

I think that Galston is mistaken. Teaching and publicising civic respect, solidarity, and tolerance cannot depend on the systematic celebration of the past infringements of liberal standards. Rather, *condemning* society's past infringements, abstaining from showing pride in them, is a fundamental component of civic education; liberal education is incompatible with the public celebration of a non-liberal history. Moreover, it is hard to believe that the *legitimacy* of liberal institutions must be defended at least partly by the promotion of gross lies.

In contrast, Rawls offers a neat reason for the adoption of a CHN for educational purposes:

If citizens of a well-ordered society are to recognize one another as free and equal, basic institutions must educate them to this conception of themselves, as well as publicly exhibit and encourage this ideal of political justice. This task of education belongs to what we may call the wide role of a political conception. In this role such a conception is part of the public political culture. (*JF*: 56. Footnote omitted)

Aside from the direct regulation of the basic structure, a conception of justice has a wider educational role amongst citizens. In such a role, a conception of justice promotes the fundamental idea that persons are free and equal. The institutional promotion of a non-critical historical narrative goes against this very role, for such a narrative is premised on the systematic omission of past infringements to the ideal of political justice.

As we have seen, adopting a CHN holds normative significance because of three interconnected reasons. First, a CHN is a way for institutions to state their condemnation of the non-liberal social context that allowed the perpetration of political violence in the past, considering that a liberal society is a cultural and political continuum over time. Thus, by promoting a CHN, institutions reaffirm their liberal integrity. In similar fashion, by demanding that their institutions promote a CHN, citizens also reaffirm their liberal integrity—for this is one of the ways in which they must discharge their duty to uphold just institutions. Second, in highly plural and multicultural societies—as it is the case of all current liberal societies—, a CHN helps to foster social stability for the right reasons. And, third, a CHN encourages a correct account of civic education.

These three reasons grant the normative significance of a CHN. Such significance anticipates a response to an objection with which I would like to finish this section. It could be argued that it should simply not be one of the tasks of the State to promote one official historical narrative over others. Rather, the objection continues, the diffusion of the different understandings

of the past should be placed completely on the hands of civil society and citizens as such. Let me call this position *the libertarian view of history*, since it advocates for leaving exclusively to citizens the creation of historical narratives within a liberal State.

I find *the libertarian view of history* implausible for at least two reasons. First, it goes against empirical evidence. *All* liberal societies we know and have known create historical narratives. Second, it seems unlikely that a liberal society could endure over time without the public dissemination of a particular historical narrative. Admittedly, the last point is an open empirical question. However, as I already said, it is not accidental that liberal States create these narratives. The goals mentioned before depend on them to a great extent. In the end, the *libertarian view of history* seems concomitant to a much more minimal understanding of the functions of the State —e.g., a State that should not provide civic education, or finance history museums, or name public places and political units, or create public monuments, or adopt days of historical commemoration, or implement an official holiday calendar, etc. Defending a liberal-egalitarian understanding of the State against this alternative minimalist framework is out of the scope of this work. Note, nevertheless, that as long as institutions assume that society has an intergenerational span, and thus use their resources to publicise a particular view of society's history, the need for adopting a CHN as opposed to a non-critical historical narrative will arise.

1.4 Concluding Remarks

This concludes my assessment of the normative significance of historical rectification. Regarding rectification of recent historical wrongs, we saw that public provisions for coming to terms with an unjust past are important for (1) citizens' psychology and self-respect and for (2) balancing citizens' moral relations to one another —that is, for ensuring their *mutual* respect as equal citizens. Finally, regarding historical rectification related to distant past wrongs, we saw that a critical historical narrative promotes (3) society's liberal integrity. In the following chapter I will address the view that, regardless of the stated importance of historical injustice in the real world, ensuring its rectification is a problem that does not arise within Rawls's theory.

2

Is Historical Injustice outside the Scope of Rawls's Theory?

In the previous chapter we saw that coming to terms with a past shaped by political violence is of the utmost importance for citizens and liberal societies in relation to (1) citizens' psychology and self-respect, (2) citizens' mutual respect, and (3) society's liberal integrity. In this chapter I will address the view that, regardless of the stated importance of historical injustice in the real world, its rectification is a problem that does not arise within Rawls's theory. By far, this is the default position in contemporary political philosophy. Both defenders and opponents of the normative significance of past wrongs for present justice tend to reiterate this idea. The *ideal* character of Rawls's theory has a powerful and misleading effect in this regard (§ 2.1). However, two other considerations may also be given in favour of this default position. These point, respectively, to the *forward-looking* (§ 2.2) and the *non-comprehensive* character of Rawls's theory (§ 2.3). In the next three sections I will show how each of these three considerations fail to establish that Rawls's theory is warranted in avoiding concern for historical injustice. Since the first of these three considerations (i.e., the ideal character of

Rawls's theory) is the one most often given to support this position, I will spend nearly half of this chapter evaluating it. I will defend two main claims in this regard. First, that victims' interest in coming to terms with a past shaped by political violence is on a par with other human psychological regularities which are already central to Rawls's *ideal* theory. Thus, I hold that such a theory should consider the psychological regularities associated with the importance of coming to terms with an unjust past just as much as it considers other psychological regularities. Second, that an *unjust pre-ordered past* is a permanent feature of a liberal well-ordered society, in the very same way as *reasonable* and *cultural pluralism* are. Thus, I hold that we as theorists need to show that even under historical conditions marked by gross injustice—as is the case of all current liberal societies—the Rawlsian utopia of a stable well-ordered society is still realistic.

2.1 The *Ideal Character* of Rawls's Theory

The first thing that comes to mind when confronted with the question of why historical injustice is not of the concern of Rawls's theory is the idea that such a task is out of the scope of what an *ideal* theory of justice is to provide. For, according to this idea, dealing with historical injustice falls exclusively under the realm of *non-ideal* theory. Even defenders of the normative significance of past wrongs for present justice share this opinion. Janna Thompson, one of the scholars that has defended most thoroughly the rectification of historical wrongs, states:

We have to drop the assumption that there is strict compliance with principles of justice in order to consider how representatives of family lines would regard claims for reparation for historical injustice. (Thompson 2001: 129)

Similarly, George Sher, by titling his book on the issue '*Approximate Justice: Studies in Non-Ideal Theory*' (Sher 1997), states his view that historical injustice is the exclusive concern of non-ideal theory. More recently, Daniel Butt, in one of the few philosophy books that addresses rectification of historical injustice at the international level, holds: 'In dealing with the rectification of historical injustice, then, this is an exercise of non-ideal theory' (Butt 2009: 6-7). There are many more examples. In fact, all of these authors are following the spirit of Rawls's first remarks on ideal theory in *A Theory of Justice*:

[F]or the most part I examine the principles that would regulate a well-ordered society. Everyone is presumed to act justly and to do his part in upholding just institutions [...] Thus I consider primarily what I call strict compliance as opposed to partial compliance theory. The latter studies the principles that govern how we are to deal with injustice. It comprises such topics as the theory of punishment, the doctrine of just war, and the justification of the various ways of opposing unjust regimes, ranging from civil disobedience and conscientious objection to militant resistance and revolution. Also included here are questions of compensatory justice and of weighing one form of institutional injustice against another. (*TJ*: 7-8)

In this passage Rawls lists compensatory justice (i.e., *specific acts of reparation* of injustices committed in the past) as part of non-ideal theory. In other places (e.g., *LP*: 89-90) he claims that transitional justice (i.e., *how* to arrive from contemporary societies to well-ordered societies) is also part of non-ideal theory. While I agree with Rawls in that both compensatory and transitional justice are out of the scope of what an ideal theory is to provide, I believe that the *importance* of coming to terms with a past shaped by gross

injustice *is* of the competence of ideal theory. To be clear, my claim is not a direct attack on the main way Rawls draws the distinction between ideal and non-ideal theory or to the importance that Rawls ascribes to ideal theory. These are the quarrels of Mills 2005, Farrelly 2007, Robeyns 2008, Boettcher 2009, Sen 2006, 2009, Mason 2010 and Wiens 2012 (attacking the importance of ideal theory), as well as of Stemplowska 2008, Swift 2008, Valentini 2009, Lawford-Smith 2010, and Simmons 2010 (defending the importance of ideal theory). Likewise, my claim should not be interpreted as a way of improving a non-ideal Rawlsian theory of justice —as Taylor 2009 and Fuller 2011 take as their task. Rather, I am suggesting that given the main way in which Rawls distinguishes the two forms of theorizing, historical injustice is also of the competence of the ideal part of a theory of liberal justice. But let me first make a few comments regarding the divide itself.

First, notice that Rawls's theory (as many others) seems to be 'ideal' in a non-technical sense (see Valentini 2009: 6 and Lawford-Smith 2010: 361): it proposes an *ideal* society at which we should aim for. This non-technical sense is captured by Rawls's famous characterisation of his theory as presenting a 'realistic utopia':

Political philosophy is realistically utopian when it extends what we ordinarily thought of as the limits of practical possibility. Our hope for the future of our society rest on the belief that the social world allows for just constitutional societies [...] The idea of this society is realistically utopian in that it depicts a social world that combines political right and justice. (*LP*: 5-6)

This non-technical sense is then complemented by a technical one: Rawls's theory is ideal so long as it works under the assumptions of *strict-compliance*

with the principles of justice and *favourable conditions* (see *TJ*: 8, 214, 308-309, *LP*: 4-10, *JF*: 13, 66, Simmons 2010: 8, Valentini 2009: 1n). Now, a crucial ambiguity with this simple definition of the technical sense comes up when it is confronted with what Rawls says about non-ideal theory: so long as non-ideal theory deals with partial compliance and non-favourable conditions, non-ideal theory is to work with cases of *injustice* (see *TJ*: 8). As the quote above from the opening passages of *Theory of Justice* states, those cases include punishment, war, opposing unjust regimes, civil disobedience, etc. What Rawls may have had in mind is that non-ideal theory is to provide guidance as *how* to proceed in those scenarios in which there is injustice. I fully agree with such an idea; however, claiming that non-ideal theory is to deal with injustice obscures two important things.

First, and most importantly, it obscures the fact that injustices can be divided into cases of *past injustice* and cases of *present injustice*. By assuming strict-compliance with the principles of justice *in the present*, ideal theory rules out that cases of *present injustice* would arise in a liberal society. However, assuming strict-compliance with the principles of justice *in the present* tells nothing about whether *past injustices* have been rectified. Second, claiming that non-ideal theory is to deal with injustice also obscures the fact that one plausible way of understanding an *ideal*-theory of liberal justice is by saying that it offers a description of what is required in order to prevent injustices from arising in a liberal society. Rawls's theory offers such a description: making a long story short, a society will be free of injustice

once its basic structure is successfully regulated by the principles of justice-as-fairness. Because of its ideal character, however, his theory does not state *how* we can arrive from our contemporary liberal societies to such an injustice-free society. That, according to the divide, is the job of non-ideal theory —see *LP*: 89-90. My claim is that, in just the same way in which ideal liberal theory offers the principles to prevent injustices from arising in a liberal society, it might need to offer a principle to prevent past injustices from remaining un-rectified. It is not the job of ideal liberal theory to tell *how* we can arrive from our contemporary liberal societies to a society in which such a principle effectively guides the institutions of the basic structure. That, according to the divide with which I have agreed, is the job of non-ideal theory.

There are two types of reasons for including concern for historical injustice as part of ideal theory. The first type springs from the importance of historical rectification at a personal level (see § 1.1). These are manifest when we attend to the reasons that Rawls offers to justify the rule of law and the importance of the coercive power of the State *even* in a well-ordered society. Here Rawls is clear that, despite the fact that his description of a well-ordered society supposes both strict-compliance with the principles of justice and favourable conditions (assumptions that, as we know, make such a description ideal), the ‘normal conditions of human life’ (*TJ*: 212) render such coercive power and an account of penal sanctions necessary. Rawls says:

It is reasonable to assume that even in a well-ordered society the coercive powers of government are to some degree necessary [...] *[e]ven under reasonable ideal conditions*, it is hard to imagine, for example, a successful income tax scheme on voluntary basis. (TJ: 211. Italics added)

The idea behind this claim is that the 'normal conditions of human life' include the fact that normal persons cannot be motivated to endorse and comply with a conception of justice exclusively through the appeal to moral considerations. Citizens are also motivated by what one might call positive stimuli (complying with the conception of justice will benefit them) and negative stimuli (not complying with the conception of justice will lead to sanctions). Rawls acknowledges that these three elements (moral reasons, positive stimuli, and negative stimuli) are the normal sources of the motivation to endorse and comply with a conception of justice. Even when Rawls is doing ideal theory, he does not list as part of the favourable conditions a tendency to go against normal psychology. Being motivated intermittently by these three "forces" is a normal condition of persons' motivation, and an ideal theory must take this into account when proposing a conception of justice:

It is clear from the preceding remarks that we need an account of penal sanctions however limited *even for ideal theory*. *Given the normal conditions of human life, some such arrangements are necessary*. (TJ: 212. Italics added)

Thus, Rawls's theory takes into account a psychological fact about human motivation in order to include in his description of an ideally just society the claim that the government is to have coercive powers to impose the directives of the principles of justice. Similarly, the evidence I have provided regarding the importance in a person's life of coming to terms with an unjust

past suggests that this is another permanent feature of normal human psychology. A person who has been the victim of political violence has a great interest in having both the State and society to recognise the gravity of the injustice that has occurred to her. My suggestion, then, is that *the interest of coming to terms with the past is on a par with other human psychological regularities similar to the ones Rawls's ideal theory is sensitive to*. I do not see any reason for Rawls to include in his theory the psychological regularities he does *but not* those related to the need of coming to terms with an unjust past.

The fact of the normal sources of motivation has dramatic results in Rawls's theory: even under the best conditions we can imagine, a liberal society will always need of a State with coercive powers. This shows how serious Rawls takes psychological facts to be in determining the content of his theory. I suggest that the psychological facts described in § 1.1 prompt similar important results: even under the best conditions we can imagine for a liberal society, citizens will always feel the need to come to terms with their past.

But some may want to dispute the analogy I have made between some salient psychological facts that Rawls considers in determining the content of his theory and the psychological facts that grant the importance of coming to terms with the past. For, indeed, Rawls explicitly leaves out some other important psychological facts of normal citizens. In particular, Rawls (*TJ*: 123-125, 465) states that envy, non-risk aversion and other 'special psychologies' are not part of the motivations of the parties in the original

position. Thus, this objection may go, just as these 'special' psychological regularities are not to motivate the parties, so do those related to the importance of coming to terms with the past.

I see two main problems with this objection. In the jargon of the moral psychology done in contemporary philosophy, it may be said that the psychological facts associated with a non-rectified past wrong create in victims (amongst other things) a form *resentment*. We can then say that the objection analogises envy to this form of resentment as equal feelings or emotions. However, Rawls (*TJ*: 467) explicitly denies that envy is a *moral* emotion.¹² Conversely, it seems clear that the resentment originated by a non-rectified past wrong *is* a moral emotion. Aaron Ben-Ze'ev 2002 offers two conditions to consider an emotion as moral: (1) whether the core of the evaluative concern of the emotion is moral and (2) whether the emotion tends to lead to beneficial moral consequences (see Ben-Ze'ev 2002: 148-149). The resentment originated by a non-rectified past wrong satisfies the two. First, the evaluative concern of this form of resentment is that one has been mistreated according to the prerogatives that justice is supposed to protect. When a violation of such prerogatives occurs, one may correctly resent its lack of restoration. Second, expression of this type of resentment may help (as we saw in § 1.2) to restore the mutual respect between victims and the rest of the community. Therefore, such an emotion, as opposed to

¹² See Ben-Ze'ev 2002 for a defence of this claim; see La Caze 2001 and Tomlin 2008 for an alternative view.

envy, is not a 'special psychology'. Rather, the stated emotion is one form of *moral* resentment that Rawls (*TJ*: 467) takes as a moral feeling.

Second, and more importantly, this objection rests on a conflation of what Rawls (*PL*: 27-28) calls the three points of view. The first one is that of the parties in the original position. The second one is that of the citizens in a well-ordered society. And the third one is ours as theorists. Because *we as theorists* think that the need to come to terms with the past is a psychological and social regularity with normative significance, we stipulate that the parties in the original position would know this. This is to ensure that, if it is the case that the society ordered by the principles selected by them does count severe injustices within its past, citizens would have the institutional means to rectify them.¹³

This takes us to the second type of reasons in favour of taking issue with historical injustice in an ideal theory of liberal justice. These reasons have to do with the social level of historical rectification (see § 1.2 and § 1.3). Another comparison may help to present them. The evidence already cited suggests that coming to terms with an unjust past is on a par with other

¹³ Tomlin 2008 has argued something similar regarding the inclusion of the psychological fact of envy as part of the *knowledge* (not as a motivation) that the parties consider in the original position. I believe that his main claim is correct, although I find support for it already in Rawls's later texts. I think that Tomlin's interpretative claims rest on a conflation of the cited three points of view (that of the parties in the original position, that of the citizens in a well-ordered society, and ours as theorist). Because *we as theorists* think that envy is not a relevant moral consideration regarding the correct election of principles of justice, we stipulate that the parties in the original position would not be *motivated* by envy. However, *pace* Tomlin 2008, this does not mean that citizens in a well-ordered society would not be prey to some form of envy sometimes and that the parties in the original position would not be aware of that fact (see *JF*: 87). In the end, as Tomlin 2008: 112 himself accepts, the parties' knowledge of this psychological fact about citizens (i.e., envy) would not alter the selection of the principles of justice-as-fairness—in particular, of the difference principle.

permanent features of liberal societies such as *reasonable pluralism* and *cultural pluralism*. Rawls's *Political Liberalism*, as part of his ideal theory, is an attempt to accommodate the former. Kymlicka's 1989, 1995, 2007 as well as Margalit & Raz's 1995 (see also Raz 1995: 170-191 and, more recently, Wall 2007) respective theories of minority rights, all of them pieces of ideal theory, are attempts to accommodate the latter. The rationale for taking these two features of liberal societies as part of ideal theory is the fact that, even by supposing strict-compliance with the principles that must regulate social cooperation and favourable conditions, the two features will not disappear.

Similarly, I claim, supposing strict-compliance *in the present* with the principles that must regulate social cooperation in liberal societies does not eliminate unjust historical interactions and the social need for assessing them morally. That is, just as religious and cultural interactions will not disappear in a well-ordered society, unjust historical interactions will not go away either. *Any* theory (be it ideal or non-ideal) of liberal justice that is sound must consider the three things when offering its principles of regulation for liberal societies. If a conception of justice does not offer a fair treatment to each liberally compatible religious or cultural group, the members of these groups can justifiably say that they are not being treated as equals. Combined, Rawls and Kymlicka's descriptions of a just liberal society ensure that this result is avoided. Likewise, a further inclusion in ideal liberal

theory might be needed so that the treatment as equals to victims of historical injustices is ensured.

Notice that by making the analogy between cultural and religious pluralism —on the one hand— and historical interactions shaped by gross injustices —on the other— as permanent features of liberal societies, I am not stating that injustices will always exist in a well-ordered society. For, since a well-ordered society is defined as a society in which no injustice occurs, this latter statement would suggest a contradiction in terms. Rather, what I claim is that *a well-ordered society will always count severe injustices as part of its pre-ordered history*. In fact, the urgency of trying to arrive at well-ordered societies is precisely the fact that right now severe injustices are taking place in our societies and terrible wrongs have occurred in their past. Once our societies are well-ordered those injustices and wrongs will be part of their history and *that will always be the case*.

Also, notice that I am not claiming that *new injustices* would arise in a well-ordered society. Statements of that sort would be evaluated in § 4.2 and § 4.3. Rather as I have said, what I claim is that the injustices committed in the pre-ordered past of a society will prevail as part of that society's history once it is well-ordered; such *past* injustices (which might go as far back as the putative origin of the society in question) will be in need of normative assessment and rectification.

But some may want to dispute the stated analogy between cultural and religious pluralism —on the one hand— and historical interactions

shaped by gross injustices —on the other— as permanent features of liberal societies. This can be done by appealing to each of the two clauses that make Rawls's theory ideal in the technical sense: the assumptions, respectively, of strict-compliance and of favourable conditions. Let me consider them in turn. Appealing to the ambiguity I have previously mentioned regarding the definition of ideal theory (i.e., that Rawls is ambiguous on whether he is supposing strict compliance only in the present or also in the past), it can be argued that Rawls has supposed all along *historical* strict-compliance. This seems to be the spirit of Thompson's passage with which I opened this section:

We have to drop the assumption that there is strict compliance with principles of justice in order to consider how representatives of family lines would regard claims for reparation for historical injustice. (Thompson 2001: 129. Italics added)

Thompson thinks that the Rawlsian clause of strict-compliance eschews concern for historical injustice within the proceedings of ideal theory. Call this interpretation of the ideal character of Rawls's theory the *historical strict-compliance reading*. If this reading of Rawls's theory is assumed, then of course historical injustices are outside of its scope. However, I would challenge this claim on four grounds.

First, consider the oddness of the reading itself. The reading states that Rawls's theory wants to offer principles to regulate the basic structure of a society that, by definition (according to the assumption of *historical strict-compliance*), has always been just. So it portrays Rawls's theoretical exercise as committed to something like the following: 'Let us assume that a society

has been well-ordered from its origin up to the present. Now I will tell you which principles will make it well-ordered from the present on'. This sounds strange enough, but there is a second theoretical problem: this reading is unable to explain the origin of a well-ordered society. It simply assumes that before the society was well-ordered, such a society did not exist, since, by definition, it never had a disordered (unjust) past. The reading thus assumes that a well-ordered society comes into existence out of nowhere, as it were.

Third, if the *historical strict-compliance reading* were correct, then Rawls's theory would be ideal *in the wrong way*. For this assumption (not having a previous disordered history) is a condition that no society in the present or the future can ever possibly meet, making Rawls's theory completely useless for the human world. This also makes it clear why the past is on a par with other permanent features of a liberal society such as cultural and reasonable pluralism: if Rawls were to assume that cultural and reasonable pluralism did not exist, this would make his theory completely irrelevant for this world, since no liberal society could ever meet such assumption. As I have said, the same thing will happen if his theory simply eschews the past.

And yet there is a fourth important ground for the rejection of the *historical strict-compliance reading*: Rawls's own emphasis on the historicity of his conception of justice. This is clearly stated in (and is one of the major themes of) Rawls's *Political Liberalism*. Rawls (*PL*: xxiv) holds that just as liberalism itself, his political conception of justice is an attempt to solve a

particular problem that has a historical origin in the 'Reformation and its aftermath, with the long controversies over religious toleration in the sixteenth and seventeenth centuries'. This makes his conception of justice only proper to the societies that have gone forward from such past. Thus Rawls states he wants to offer a conception of justice for *our* liberal societies, for societies having *our* shared political culture. But consider what makes these societies *ours*. Rawls himself appeals to their past. It is because of the historical interactions that allowed the free exercise of practical reason that now our societies have the political culture they do. However, such historical interactions are replete with violations to the principles that are to regulate social cooperation as well (that is, replete with injustices). In fact, it is only because of this that the main attempt to solve the problem of political liberalism ('what are the fair terms of cooperation between citizens characterised as free and equal yet divide by profound doctrinal conflict?' — *PL*: xxv—) is so pressing.

Summing up, the *historical strict-compliance reading* of the ideal character of Rawls's theory is completely inadequate since: first, it makes Rawls's theoretical exercise odd (i.e., offering a conception of justice for a society that is assumed to be just ever since its origin); second, it is unable to explain the origin of a well-ordered society (for such society is assumed not to have a previous disordered past); third, it makes Rawls's theory ideal in the wrong way (for no liberal society in the present or the future could meet

such an assumption); and, fourth, it flies in the face of some of the most salient theses of Rawls's political liberalism.

So let me turn to the second way in which the analogy between cultural and reasonable pluralism —on the one hand— and an unjust past —on the other— as permanent features of liberal societies may be disputed. It may be argued that Rawls's clause stipulating *favourable conditions* grants the supposition that no major historical injustice has occurred in a society that is to be well-ordered at a given time by the principles of justice-as-fairness. Korsgaard best expresses this putative interpretative claim:

Ideal theory is worked out under certain assumptions. One is strict compliance: it is assumed that everyone *will* [i.e., in the present and future] act justly. The other, a little harder to specify, is that *historical*, economic, and natural conditions are such that realization of the ideal is feasible [...] We also assume in ideal theory that *there are no massive historic injustices, such as the oppression of blacks and women, to be corrected*. The point is to work out our ideal view of justice on the assumption that people, nature, and *history* will behave themselves so that the ideal can be realized, and then to determine — in light of that ideal— what is to be done in actual circumstances when they do not. (Korsgaard 1996a: 147-148. Italics added)

Korsgaard thinks that the Rawlsian clause of favourable conditions eschews concern for historical injustice within the proceedings of ideal theory. Call this interpretation of the ideal character of Rawls's theory the *favourable historical conditions reading*. If this reading of Rawls's theory is assumed, then of course historical injustices are outside of its scope. However, I would challenge this claim on two grounds.

First, notice that it is a matter of interpretative dispute whether the clause of favourable conditions is in place to grant the assumption that no major historical injustice has occurred in a society that is to be governed by

justice-as-fairness's principles. Certainly, some passages of Rawls's theoretical corpus may seem to suggest that (see, in particular, *JF*: 64-65). However, I believe that for the most part the favourable conditions clause is concerned with ensuring that, *once a well-ordered society is established*, contingent historical, natural, and social conditions from then on would favour the stable continuation of such society from one generation to the next (see, in particular, *JF*: 4, 13, *PL*: 336).¹⁴ It should also be noticed that when Rawls mentions *historical contingencies* he usually refers to the specific contingent social positions that would bias the selection of principles of justice in the original position (social class, gender, race, sexual orientation) and not to past historical interactions as such (see *PL*: 271-274 and Simmons 2010: 13-15). So, at the very least it should be accepted that the Rawlsian theoretical corpus allows a different interpretation of its ideal character than the one defended by the *favourable historical conditions reading*.

Moreover, other Rawlsian passages suggest the complete rejection of the stated reading. Consider the following:

In addition, the veil of ignorance (§24) is interpreted to mean not only that the parties have no knowledge of their particular aims and ends (except what is contain in the thin theory of the good), but also that the *historical record is closed to them*. They do not know, and cannot enumerate the social circumstances in which they may find themselves. (*TJ*: 160. Italics added)

Here Rawls states that the course of history is closed to the parties in the original position. Another passage states the same claim:

¹⁴ James 2005: 282 offers independent support for this reading of the favourable condition clause. According to him, Rawls's method for proposing principles of justice for a given practice requires us to: 'identify an existing social practice, including its point, or the goods that is meant to realize. Assume circumstances favorable *to its continuance*' (Italics added).

Let us distinguish between three kinds of facts: the first principles of social theory [...] general facts about society [...] and finally, particular facts about individuals [...] In the original position the only particular facts known to the parties are those that can be inferred by the circumstances of justice. While they know the first principles of social theory, *the course of history is closed to them*; they have no information about how often society has taken this or that form, or which kinds of society presently exist. (TJ: 175. Italics added)

Rawls reiterates that knowledge about the history of their society is closed to the parties. By doing this, Rawls has stipulated that the parties in the original position cannot make assumptions regarding the course of the past of their society. Thus, just as the parties do not know whether they or the citizens they represent would adopt this or that religion once a well-ordered society is established, *they do not know whether their well-ordered society would count severe historical injustices as part of its pre-ordered history*. In fact, this is a very good reason for the parties to ensure provisions for coming to terms with the past: for all they know, it *may be the case* that their well-ordered society counts historical injustices as part of its past.

Second, I do not think that, as Korsgaard claims, the ideal (in the non-technical sense) of a well-ordered society cannot be realised unless we suppose that history has 'behaved' itself. Rather, just as we as *theorists* need to show that it is still a practical possibility that a well-ordered society would be stable over time given the fact of persons' normal sources of motivation and the facts of reasonable and cultural pluralism, so we also need to show that it is still a practical possibility that such a society would be stable even if it counts severe injustices as part of its past. For all of what Korsgaard has stated, we could also have supposed that nature has 'behaved' in a way that

citizens do not feel a normal tendency to take advantage of social cooperation sometimes. The putative realisation of the ideal of a well-ordered society would be easier under this supposition. However, it is a remarkable virtue of Rawls's theory that it does not eschew normal psychology and thus confronts the challenge of showing that the ideal of a well-ordered society is possible for normal human beings. Similarly, we need to show that a past marked by injustice does not hinder the possibility of arriving at that ideal.

To state my point more clearly, recall what Rawls (*PL*: 27-28) calls the three points of view: the one of the parties in the original position, the one of the citizens in a well-ordered society, and the one of us as theorists. The question then is this: from which point of view are we to suppose that history has 'behaved' itself so that no historical injustice has occurred? As we just saw, it cannot be that of the parties of the original position, for the passages just quoted exclude this possibility. It cannot be that of citizens of liberal societies because the previous history of their well-ordered society, just as other psychological and social facts, is something that simply happens to them. And it cannot be ours as theorists, for we actually know that the history of liberal societies is full of injustices and *that*, I tend to believe, is one of our main motivations to discuss how a well-ordered society would look like.¹⁵

¹⁵ Consider what Rawls (*JF*: 13) holds: 'We ask in effect what a perfectly just, or nearly just, constitutional regime may be, and whether it may come about and be made stable under the circumstances of justice, and so *under realistic though reasonably favorable conditions*. In this way justice as fairness is realistically utopian: it probes the limits of the realistically practicable, that is, how far in our world (given its laws and *tendencies*) a democratic regime can attain complete realization of its appropriate political values' (Italics added). Considering what Rawls states here, another way to put my second objection to *the favourable historical conditions reading* is that it holds an *unreasonable* assumption, for it is unrealistic and also

Summing up, the *favourable historical conditions reading* of the ideal character of Rawls's theory is both interpretatively and theoretically incorrect. First, since the course of history is closed to the parties in the original position, the need for ensuring provisions for coming to terms with an unjust past is one consideration that they must ponder (just as they must ponder the psychological fact of citizens' sources of motivation, and the social facts of reasonable and cultural pluralism). Second, we as theorist need to show that even under historical conditions marked by injustice —as is the case of all current liberal societies— the Rawlsian utopia (a stable well-ordered society) is still realistic.

For all of the aforementioned reasons, the phenomenon of the importance of coming to terms with an unjust past must be conceptualised also as part of ideal theory. Both the personal and social levels of it grant such conceptualisation. As we have seen, this phenomenon is on a par with other psychological regularities which are already central to Rawls's ideal theory. Likewise, an unjust past and the need to assess it morally is on a par with other permanent features of liberal societies such as cultural and reasonable pluralism, features that Rawls's ideal theory considers quite carefully.

Remember that I am not asking ideal theory to flesh out *how* we can arrive from our contemporary liberal societies (in which most relevant historical injustices remain un-rectified) to well-ordered societies (in which no relevant historical injustice remains un-rectified). Theories of transition

unnecessary to claim that history has 'behaved' itself (as Korsgaard says in the passage quoted) in order to show that the ideal of a well-ordered society can be realised.

may thus remain within the scope of non-ideal theory. Also, I have accepted that compensatory justice (stating *specific acts of reparation* for injustices committed in the past) is beyond the scope of what an ideal theory is to provide. What I am asking is for ideal theory to be a guide for non-ideal theory (as defenders of the divide claim it must be —see *JF*: 66, Buchanan 2004: 60, Valentini 2009, Swift 2008, Simmons 2010) in cases of past injustice. In this sense, Rawls states:

So conceived, nonideal theory presupposes that ideal theory is on hand. For until the ideal is identified, at least in outline —and that is all we should expect— non-ideal theory lacks an objective, an aim, by reference to which its queries can be answered. (*LP*: 89-91)

In line with this characterisation, my claim is Rawls's theory should also offer 'an aim by reference to which' non-ideal theory's queries regarding rectification of historical injustice can be answered.

Finally, let me finish this section by making the following remark. I have analogised reasonable pluralism, cultural pluralism, and unjust pre-ordered interactions as permanent features of well-ordered societies. In passing, I have mentioned that Kymlicka's and similar theories of minority rights might ensure equal treatment to members of minority cultures (just as Rawls's political liberalism might ensure equal treatment to reasonable comprehensive doctrines). Regardless of whether this claim is correct (which is indeed irrelevant for the purposes of this thesis), let me underline a structural similarity between Kymlicka's critique of Rawls's theory and mine.

Kymlicka has consistently defended two claims, one empirical in character and another normative. First, that no liberal society is a culturally

homogeneous political community (see Kymlicka 1989: 206-220, 1995: 10-26). As I mentioned in § 1.3, all liberal societies, aside from observing pluralism of comprehensive doctrines, observe strong forms of cultural pluralism: they contain national minorities (such as Scotland within the United Kingdom), indigenous communities (such as the Inuit in Canada), and immigrant populations (such as the vast Mexican community in the United States). Second, Kymlicka holds that Rawls's theory of distributive justice as it is written does not have the normative resources to ensure a just treatment to such cultural minorities. Therefore, Kymlicka argues, including a scheme of minority rights (commanding a just treatment to cultural minorities) within such a theory should fill this theoretical deficit (see Kymlicka 1989: 162-205, 1995: 75-130).

Similarly, throughout this work I defend two claims, one empirical in character and another normative. First, that every single liberal society counts gross injustices within its past, and this would remain the case once they became well-ordered (§ 1.2-2.1). Second, that Rawls's theory as it stands right now might not have the normative resources to ensure a just treatment to victims of such historical wrongs (§ 3.1).

Now consider most critiques of Kymlicka's second claim: they focus on whether such a putative scheme of minority rights conflicts with the original precepts of justice-as-fairness such as the importance of equal liberties and fair-equality of opportunity (see, paradigmatically, Barry 2001). No one, to my knowledge, has criticised Kymlicka's second claim on the basis that

Rawls's theory is ideal and, therefore, justified in supposing —against all empirical evidence regarding the contemporary world— that to every liberal society corresponds a culturally homogeneous community. No one has criticised Kymlicka's second claim on that basis despite the fact that Rawls is explicit about the stated supposition:

Historical conquests and immigration have caused the intermingling of groups with different cultures and historical memories who now reside within the territory of most contemporary democratic governments. Notwithstanding, the Law of Peoples starts with the need for common sympathies, no matter what their source may be. My hope is that if we begin in this simplified way, we can work out political principles that will, in due course, enable to deal with more difficult cases where all the citizens are not united by a common language and share historical memories. (*LP*: 24-25)

Contrary to the case of unjust historical interactions, Rawls is explicit about the fact that his theory would simply assume cultural homogeneity within each political society. If we frame this as one of the suppositions granted by the two stipulations that make Rawls theory ideal, we could say that the *favourable conditions* clause grants the supposition. But if this is so, why is it that no one has criticised Kymlicka's second claim on this basis? The answer, it seems to me, is that simply appealing to the ideal character of Rawls's theory would be an inadequate response to Kymlicka's second claim. In similar fashion, appealing to the ideal character of Rawls's theory is an inadequate response to my second claim (i.e., that Rawls's theory as it stands right now might not have the normative resources to ensure a just treatment to victims of historical wrongs). Rather, a proper response should argue that the principle I will explore in § 3.2 to fill this gap in Rawls's theory is either unnecessary to ensure historical rectification or, as opponents of Kymlicka's

second claim vigorously defend, in conflict with the fundamental precepts of justice-as-fairness. In any event, pointing to the ideal character of Rawls's theory is simply not the correct way of dealing with normative significance of historical injustice rectification.¹⁶

2.2 The *Forward-Looking Character* of Rawls's Theory

Once we have seen that appealing to the *ideal* character of Rawls's theory does not justify excluding concern for historical injustice from his theory, I now turn to a second putatively Rawlsian consideration stating that the description of a just society is warranted in being strictly forward-looking.

Something similar to this consideration is expressed by John Simmons:

Rawls's ideal theory, it is said, pays no attention to the long histories of injustice [...] But the kind of "oversights" with which such criticism charge Rawls are not in any way essential to Rawls's characterisation of ideal theory or its relationship to non-ideal theory; they are rather simply a function of the specific approach that Rawls employs to derive the content of ideal theory, to argue for his particular, favored conception of "justice as fairness". Historical injustice goes unaddressed in Rawls's theory because the derived principles of justice are purely "forward-looking", *because the choice problem given to Rawls's original position contractors requires their choice of forward-looking principles*. (Simmons 2010: 32-33. Italics added)

Unlike Korsgaard 1996a: 147-148, Sher 1997, Thompson 2001: 129, Butt 2009: 6-7, and many others, Simmons does not believe that the *ideal character* of Rawls's theory excludes concern for historical injustice. Rather, Simmons thinks that the choice problem presented to the parties in the original position imposes the selection of purely forward-looking principles. I believe Simmons is mistaken. For, as we have seen, Rawls is quite clear in

¹⁶ In fact, in § 4.2 I will defend that the ideal theory reply is equally inadequate to answer similar critiques to justice-as-fairness regarding its lack of concern for race and gender oppression.

stating that the parties in the original position must know all relevant psychological and social facts regarding human nature and social regularities:

It is taken for granted, however, that they [the parties] know the general facts about human society. They understand political affairs and the principles of economic theory; they know the basis of social organization and the laws of human psychology. Indeed, the parties are presumed to know whatever general facts affect the choice of principles of justice. (*TJ*: 119)

The psychological and social evidence cited in § 1.1-1.3 seems sufficient for stating that coming to terms with the past is of the greatest relevance for human beings and liberal societies. As such, this phenomenon must be one of the facts available to the parties in the original position, since they do not know whether their resultant well-ordered society would count severe historical injustices as part of its past. Because of this simple reason, the choice problem presented to the parties in the original position (i.e., selecting principles to regulate social cooperation under the veil of ignorance in full knowledge of all relevant information regarding the regularities of human psychology and society —see *TJ*: 102-160) cannot be correctly solved by adopting strictly forward-looking principles.

It may be that behind Simmons's verdict of what warrants the forward-looking character of Rawls's theory lies a reading of the original position more akin to rational choice theory, for rational choice theory tends to work under the assumption that persons behave in a forward-looking way. However (as it has been neatly exposed by Sen 1977, Herzog 2006: 35ff, and many others), there are several arguments showing that such assumption in

rational choice theory is unwarranted. For my part, I find it sufficient to reject the stated assumption the fact that it goes against human normal psychology. As we saw in § 1.1, the understanding of the past plays a major role in persons' own identity formation and choices. More importantly for the present purposes is the fact that Rawls was quite clear in his later work regarding his dissatisfaction with the rational choice reading of the original position (see *JF*: 82n).

Moreover, Rawls (*TJ*: 260) already offers a neat reason against Simmons's understanding of the choice problem given to the parties in the original position. Rawls's justification of the principle of just savings depends, precisely, on denying that *time preference* is one of the concerns of the parties in the original position:

The original position is so defined that it leads to the correct principles in this respect. In the case of the individual, pure time preference is irrational; it means that he is not viewing all moments as equally parts of one life. In the case of society, pure time preference is unjust: it means (*in the more common instance* when future is discounted) that the living take advantage of their position in time to favor their own interests. (*TJ*: 260. Italics added)

Here Rawls states that time preference in the case of society is unjust. Of course, he adds that its most common instance is to show no concern for future generations or for the future of the current generation. However, that is only one instance of the injustice produced by time preference. It seems that the only other instance would be not to show concern for previous generations or for the past of the current generation. It thus must be accepted that the very same reason that is behind the justification of the principle of just savings compels the Rawlsian theory to show concern for

historical injustices. If that reason is introduced in the justification of one of the principles of justice, then the forward-looking character of the theory — just as Rawls's forward-looking conception of the person, as we saw in § 1.1 — must be abandoned.

One final reason against the exclusive forward-looking character of Rawls's theory springs from what Rawls calls the 'most fundamental idea' of his conception of justice: i.e., that a liberal society must be understood as a 'fair system of social cooperation over time from one generation to the next' (*JF*: 5). Rawls repeatedly comes back to the idea of a temporal continuum behind his understanding of society: 'Recall that a political society is always regarded as a schema of cooperation over time indefinitely' (*JF*: 162). However, if society is so understood by a theory, such a theory must offer guidance with regard to several problems arising between generations that interact with each other. Rawls takes issue with a relevant one having to do with a present and a future generation (the problem of just savings). But just as such a problem arises by assuming that society is a temporal continuum, the social dimension of the importance of coming to terms with the past arises too by the very same assumption. This last consideration is very important, for even if the rationale given in favour of the just savings principle was to be abandoned (as English 1977: 98, Wall 2003: 79, Nagel and Parfit —see *JF*: 160n— urge Rawls to proceed), this would not change the fact that Rawls's theory is required to give equal treatment to the problems produced by the interaction of a present generation with both its

past and future generations. In this sense, Thompson is mistaken when she suggests:

John Rawls, for example, thinks of justice between generations as consisting of duties that citizens owe to their descendants. This conception of justice divides justice into two parts: synchronic and diachronic. *Synchronic justice* is justice between contemporaries, or between those contemporaries who are full participants in the political relationships of the society. *Diachronic justice* has to do with relationships between these contemporaries and future citizens. (Thompson 2009: 2. Italics in the original, references removed)

Thompson 2009: 2ff seems to suggest that the stated dichotomy is inadequate because the problems of interacting generations are also *synchronic* problems: that is, problems of 'justice between contemporaries'. That this is so is manifest when we consider how the duties of society with further generations and to rectify historical injustice have effects in the distribution of primary goods amongst contemporaries. Consider what Butt claims when analysing the impact of rectificatory justice on distributive justice:

Questions of rectificatory justice in response to historic wrongdoing are questions which concern the distribution of burdens and benefits in the present day. Claims that a given group is owed compensation as a result of [a] historic injustice [...] are claims about who should have what here and now. They are claims that operate in the real world, and that propose particular courses of action which affect the distribution of resources within, and between, societies. (Butt 2009: 33)

Butt points at the fact that the decision of how many resources will be allocated for the task of rectifying historical injustices will have an impact on how the resources amongst contemporaries are to be distributed. We can find a similar claim in Rawls's characterisation of the problem of just savings, since any solution contemporaries give to this problem will affect the distribution of primary goods amongst them: if they decide to leave as many

resources to the next generation as they have, this will set a limit to their use of resources for the production and distribution of primary goods.

This makes it clear that for Rawls the problem of just savings is also a *synchronic* problem, and thus Thompson's 2009: 2 conceptualisation of Rawls's theory as dividing justice into *synchronic* and *diachronic* is incorrect. Moreover, as we have seen, at the very basis of Rawls's understanding of how a liberal society must be conceived of there is a commitment to the idea that 'intergenerational relationships are central to a political society' (Thompson 2009: 12), for this is implied by Rawls's insistence on the idea that a society is a continuum in time. Interactions between generations are intrinsic to this understanding of society. What Rawls missed, perhaps due to the ambiguity I have pointed out in the previous section regarding his characterisation of the divide between ideal and non-ideal theory, is that some of the problems arising from those interactions, which are relevant for his own theory of justice, have to do with the relation of a current generation with its own past or with previous generations.

So we have three reasons against the forward-looking character of Rawls's theory. First, the parties in the original position know all relevant psychological and social facts, one of which must be the importance of coming to terms with the past for normal human beings and liberal societies, especially considering that they do not know whether their resultant well-ordered society would count severe injustices as part of its pre-ordered history. Second, since time preference is ruled out by the construction of the

original position, the parties are to show concern for generations behind them (just as they are to show concern for generations next to them). And third, even if this rationale for the just savings principle were to be completely abandoned, so long as Rawls assumes that a liberal society is a continuum in time, his theory is still to show concern for the problems of interacting generations. These problems include not only the interaction with future generations but also with past generations. Considering these three reasons, *pace* Simmons 2010, the choice problem modelled in the original position does not warrant the absence of concern for historical injustice.

However, it could still be argued that Rawls's theory, as any other theory of liberal justice, is justified in being exclusively forward-looking because of the putative existence of the so-called *non-identity problem*. Typically, such a problem is framed as follows: had the past injustice to victims not occurred, victims' descendants simply would not have come into existence. Because existing (at least in sufficiently bearable conditions) is better than failing to exist, the occurrence of such past injustice does not constitute a harm to descendants. Thus, no compensation, apology, or any other rectificatory measure is due on descendants' behalf for what in fact made possible their very existence (see Sher 1981: 7-8, 2005: 181-182, Morris 1984: 177ff, Waldron 1992: 12, Simmons 1995:178n, A. Cohen 2009: 81-83).

All versions of the *non-identity problem* in the literature begin by accepting the following standard modal thesis (call it the *necessity of origin*

thesis): it is a necessary condition of a person's existence to be originated by a particular pair of gametes. In possible-worlds talk, this is equivalent to say that there is no world in which a person exists and a different pair of gametes originates her.¹⁷ From here, we can distinguish two different forms in which the non-identity problem is presented: a modal and a probabilistic version (see Sher 2005: 184-185).

The *modal version* states a strong claim about modal identity: i.e., that the past injustice is a *necessary* condition of the existence of victims' descendants and because of that it cannot constitute a harm to them. Had such injustice not occurred, their parents (or grandparents, or great-grandparents, and so on) would not have met and conceived them (or conceived their parents, or grandparents, and so on) at the time they did, which is biologically required in order for them to be originated by the same pair of gametes. In possible-worlds talk, this is equivalent to saying that there is no world in which the original harm did not occur and descendants exist.

Before stating my own response, it is worth mentioning two recurrent responses to the *modal version*. First, consider the *further-harm response*. It can be argued that even if the past injustice is a necessary condition of descendants' existence, rectification is still due on the following basis: while the original injustice (call it Injustice 1) harms the members of the generation that suffered such an injustice (call it Generation 1), the lack of rectification

¹⁷ Perhaps this thesis is most famous because of Saul Kripke's formulation (see Kripke 1980: 110ff). Formally, such a thesis is an instance of the following: for all x and for all y , if x originates from y , then x originates from y in all worlds in which x exists.

of Injustice 1 to Generation 1 creates a *further* harm (call it Injustice 2) to the next generation of descendants (call it Generation 2). Similarly, the lack of rectification of Injustice 2 to Generation 2 creates a further harm (call it Injustice 3) to the next generation of descendants (call it Generation 3), and so on. Relevantly, it is not the case that Injustice 1 harms Generation 2. Rather, the fact that Injustice 1 was never rectified for Generation 1 brings about Injustice 2, which harms Generation 2 and thus is the one for which Generation 2 is owed rectification. Likewise, the fact that Injustice 2 was never rectified for Generation 2 brings about Injustice 3, which is the one harming Generation 3 and thus the one for which Generation 3 is owed rectification. In this way, a non-rectified original injustice creates a chain of further injustices harming successive generations of descendants (see Boxill 2003: 89, Sher 2005: 190-195, and Butt 2009: 188-189 for independent defences of this response).

Second, consider the *group-harm response* against the *modal version*. It could be argued that even if the past injustice is a necessary condition of descendants' existence, and because of that it does not harm any particular descendant taken as an individual, such injustice nevertheless harms all descendants because of their belonging to the harmed group (see Herstein 2008: 527-531 for a defence of this response).

While I find both of these responses plausible, they have complications. For instance, one problem of the *further-harm response* is that it does not fit the way in which public official apologies regarding historical

injustices are given. For these apologies are not, or not only, offered due to the further harms that a non-rectified original historical injustice creates, but also because of the occurrence of the *very* original historical injustice. That is, a public apology is due not only because of the fact that the original injustice has not been rectified over the course of history but also because of the fact of its very occurrence (see Herstein 2008: 519-523). Likewise, the *collective-harm response* is committed to the idea that there might be some injustices harming a collective even when *benefiting* each particular individual forming the collective —so long as the harm brings each of them into existence— (see Butt 2009: 105).

Of course, I am not stating that these and similar objections cannot be overcome by the responses just explored. Yet, another response seems to be more straightforward. Let me repeat the main claim of the *modal version* using possible-worlds talk: there is no world in which the original harm did not occur and descendants exist. This claim about descendants' modal identity, unlike the standard *necessity of origin thesis*, is far from being intuitively correct. For it is easy to think of several worlds in which descendants exist without the original harm made to their ancestors. Take the case of the enslavement of Africans and the existence of actual African-Americans. History could have been incredibly different from the way it was *and* still leave room for African-Americans' actual ancestors to meet. For one, they could have been invited to go to North America as workers with equal rights and moral status. In possible-worlds talk, this is equivalent to saying

that there is a world identical to the actual one (relevantly, containing the same individuals) except for the fact that Africans were invited to go to North America as workers with equal rights and moral status. This seems to be a genuine possibility regarding descendants' modal identity (see Simmons 1995: 174n and Butt 2009: 115 for equivalent responses).¹⁸

Derek Parfit himself, who is typically credited with the introduction of the non-identity problem (see Parfit 1984: 351-355, 2011: 218), refrains from endorsing the controversial necessity claim of the *modal version* when he states:

(TD2) If any particular person had not been conceived within the month of the time when she was in fact conceived, he would in fact never have existed.

I claim that [what TD2 holds] is *in fact* true. I do not claim that it is *necessarily true*. The different views about this subject make competing claims about what is necessary. It is because I claim less that my claim is not controversial. Those who disagree about what *could* have happened may agree about what would *in fact* have happened. (Parfit 1984: 351 Italics in the original)

Here Parfit clarifies that he does not hold that the *actual* course of history, affecting monthly patterns of conception, is a *necessary* condition of actual persons' existence. Rather, his (TD2) is compatible with statements like the following: it *could* have been the case that history was different (and thus monthly patterns of conception had variations) and actual persons still existed. Coming back to the example I just introduced, Parfit's original formulation of the problem is compatible with the claim that it *could* have

¹⁸ Butt states: 'to assess harm following injustice, it is necessary to compare the current day injustice with some kind of counterfactual. This is done by imagining a possible world in which no injustice occurred. However, *there are many such possible worlds, as there are many possible kinds of interaction between the victim and the offender*. One possible world is the world where the act of injustice simply did not take place' (Butt 2009: 115. Italics added).

been the case that African Americans' ancestors were invited to North America as workers with equal rights and moral status and their actual descendants still existed. In contrast, as we have seen, the *modal version* is committed to an extremely polemical assumption about persons' modal identity. Until a defence of such an assumption is provided, this version can be set aside.¹⁹

So let me turn to the *probabilistic version* of the non-identity problem. This version makes the following softer claim: it is highly *unlikely* that victims' descendants would have existed had the past injustice against their ancestors not occurred. Indeed, Parfit's own formulation is closer to the *probabilistic version*. He states:

Which particular children we have depends on the slightest details of our private lives. Many of our acts affect such details in our own and other people's lives, and these effects spread, like ripples in a pool, over more and more lives. Unlike ripples, moreover, these effects never fade away. Over time, there will be more and more people of whom it is true that, if we acted differently, those people would never have been conceived. *If the motor car had not been invented, for example, it is likely that readers of this book would never have existed.* (Parfit 2011: 218. Italics added)

As the last statement of this passage suggests, Parfit seems to have intended to pose the *probabilistic version* (rather than the *modal version*) of the non-

¹⁹ This also shows that Sher's original response to the non-identity problem, based on what he called the *branching criterion of possible worlds* is completely inadequate (see Sher 1979: 381-382 and more recently, Sher 2005: 187). This criterion is committed to an even stronger claim about descendants' modal identity than the *modal version*. For according to such a criterion *actual* history as a whole prior to my conception is a *necessary* condition of my existence (see Sher 1979: 382, 2005: 187). In possible worlds talk, this is equivalent to stating that there is no world in which history prior to my conception is different from *actual* history and I exist. Yet another way of stating the same: I could not have existed had history prior to my conception been slightly different from the way it actually was. However, this form of hyper-essentialism states an extremely polemical thesis about modal identity. For it follows that statements of the sort 'I could have been born in the UK had my parents been studying there in 1982' are plainly false. Even worse, if Sher's *branching criterion of possible worlds* is correct, then statement of the sort 'I could have existed even if Julio Cesar had not died in the Roman Senate but in his house in the year 44 BC' are false.

identity problem. Parfit's disclaimer regarding the non-necessary character of his (TD2) —stated in the passage previously quoted— also reinforces this interpretation.

Now, the *probabilistic version* (unlike the *modal version*) seems to state a fairly uncontroversial claim. It seems indeed highly unlikely that descendants would have existed had the original historical harm to their ancestors not occurred. For, coming back to our example, even when it is a genuine modal possibility that actual African Americans could have existed had their ancestors been invited to North America as workers with equal rights and moral status, this would have required of the occurrence of an extremely unlikely chain of events stretching on for centuries resembling the actual one.

Nevertheless, I think that the relevant question to ask is what the moral significance of this fairly uncontroversial probabilistic claim is. For if we ascribe to it the significance of justifying the avoidance of historical rectification, then we are committed to the idea that any meaningful theory for the moral assessment of the past should accept absurd conclusions. To show this, let me introduce the following example.

Suppose that my parents travelled from very far distant cities to the capital in order to join an activist group against the Vietnam War and only because of this reason they met each other. Considering this, it is indeed very *unlikely* that I would have existed had the Vietnam War not occurred — assuming, as we are, that my parents travelled to the capital for this reason

and otherwise would have pursued their independent lives in their respective cities. Does this mean that I cannot morally object to the occurrence of the Vietnam War without thereby stating that I would like not to have existed? Or even worse, should I be grateful about the occurrence of the war, so long as I am grateful about my own existence? These inferences, drawn by the *probabilistic version*, seem absurd.

More importantly, the *probabilistic version* simply fails to show the moral significance of its uncontroversial probabilistic claim. For, so long as it is a genuine modal possibility that I could have existed had the war not occurred—however remote such a possibility is—, this fact about my modal identity allows me to morally object to the war without thereby objecting to my very existence. Similar claims can be made about historical wrongs without the occurrence of which it would have been very unlikely that actual descendants had existed (see Simmons 1995: 174n and Butt 2009: 106 for similar conclusions).

Because of these reasons, neither the *modal* nor the *probabilistic* versions of the non-identity problem succeed in showing that rectification of past wrongs is not morally required, and thus in showing that a liberal theory of justice such as Rawls's is warranted in being strictly forward-looking. As we have seen, while the *modal version* rests on an undefended and highly polemical necessity claim, the *probabilistic version* fails to show the moral significance of the non-controversial probabilistic claim on which it is based.

2.3 The *Non-Comprehensive Character* of Rawls's Theory

So far we have seen that the *ideal* and *forward-looking* character of Rawls's theory does not justify the absence of concern for historical injustice in Rawls's description of a well-ordered society. In this final section I will explore a third Rawlsian consideration that could be offered in favour of not conceptualising of historical rectification as part of such a theory: i.e., the claim that, in so doing, Rawls's theory would acquire a *comprehensive character* that does not suit the purposes of a political conception of justice, considering the fact of reasonable pluralism. In fact, I do believe that some approaches to the moral assessment of the past are comprehensive in character. Three concerns in this direction occur to me.

The first and clearest way in which an account pointing at the importance of historical injustice becomes comprehensive is if such an account offers an ethics of memory or remembrance for the personal domain. This is the case of Avishai Margalit 2002 and Jeffrey Blustein 2008, for these authors provide (amongst other things) an account of what persons *ought* to do regarding the understanding of their own past. Margalit states:

My question, 'Is there an ethics of memory?', is both about microethics (the ethics of individuals) and about macroethics (the ethics of collectives). What I want to address can be rendered by a series of questions: Are we obligated to remember people and events from the past? [...] I reach the conclusion that while there is an ethics of memory, there is very little morality of memory. (Margalit 2002: 6-7)

Margalit distinguishes ethics from morality in the following sense: while morality is concerned with the principles and values that must guide our relations with all human beings as human beings, ethics is restricted only to

the principles and values that must guide our relations with our proximate human beings (friends, family, political community). Considering this distinction, when Margalit states that there is an *ethics* of memory, he is proposing that we *ought* to behave in certain ways in a domain that includes our relationships with friends and family.²⁰

Blustein also offers an ethics of memory for personal relationships. He holds that by the very fact of remembering an event a person takes responsibility for it. Accordingly, remembrance is connected with moral responsibility: one must assume responsibility for what one has done (Blustein 2008: 33). However, Blustein thinks that a person must also remember certain things for which she is not responsible in any sense. For instance, Blustein believes that persons have an ethical (in Margalit's sense) obligation to retain the memory of their dead parents and close friends —or, as Blustein 2008: 245 calls them, 'the dear departed'.²¹

The personal domain of these respective ethics of memory is out of the scope of what a non-comprehensive account of historical rectification can provide. I am not claiming that it is the function of a liberal State to promote any moral obligation in this regard. So long as their distinctive accounts are

²⁰ These *oughts*, according to Margalit, are of the following conditional form: *if you want to have correct ethical relationships, you must remember certain things about the persons you are engaging with* (see Margalit 2002: 104-106).

²¹ Blustein's rationale for this claim is threefold: first, remembering the dear departed is a way of retaining the significance of their life (see Blustein 2008: 260-263); second, by remembering the dear departed we respect the 'enduring duties' of love and honour on which our moral relationship to them is based (see Blustein 2008: 273-276); and, third, so long as we want to be remembered in turn by other persons, we shall remember the dear departed following an impulse of reciprocity (see Blustein 2008: 276-281). While the first two create unconditional duties of remembrance, the third one only creates a conditional duty.

not the only ethical frameworks available in this regard (remember Strawson 2004 —end of § 1.1) and, more importantly, are not part of the political culture of liberal societies, Margalit and Blustein's approaches to the moral assessment of the past are comprehensive. In contrast, notice that my argument is rather different: considering the psychological evidence provided, it seems to be a human psychological regularity that *most* normal victimized citizens are in need of coming to terms with their *public* past. It is this interest that must be accounted for in a theory of liberal justice.

A second way of comprehensiveness comes by affirming highly speculative metaphysical claims about persons and the importance of their past. An instructive way of dealing with this question is provided by Bruce Ackerman's paper 'Temporal Horizons of Justice' (Ackerman 1997). Ackerman develops a series of ideas regarding 'the metaphysical foundations of the self's encounter with time', foundations in which, Ackerman argues, 'liberals may find it necessary to probe more deeply' (Ackerman 1997: 317). Ackerman's adventure into the metaphysics of the self ends up suggesting his 'own strong priority to the self's struggle for a meaningful autobiography over the claims of particular relationships and life experiences' (Ackerman 1997: 317). In contrast, notice that I am not making a metaphysical claim regarding the relationship between self and time. The relevant evidence that warrants the concern with historical injustice and the past is either empirical in character or independent of polemical speculative claims. In fact, this idea

guides the evidence I provided in § 1.1-1.3 in favour of the normative significance of historical injustice rectification.

David Velleman also offers (although indirectly) a comprehensive view about the relevance of the past. Velleman holds that the *self* narrates its own life, bringing certain past facts to light and eschewing others, creating a single coherent narrative of a life. The result, according to Velleman, is that the self invents itself. The main focus of his discussion is whether such a *narrative self* is fictional (as Daniel Dennett 1989, 1992 holds) or not. For Dennett, it is my brain the author of my self-narrative (not me or *myself*) and the result of such invention is plain fiction. For Velleman, *myself* is the author of my self-narrative and the result of such invention is true:

I shall not be arguing against [Dennett's] positive conception of the self as the fictive protagonist of a person's autobiography. On the contrary, I'll argue that Dennett's positive conception of the self is largely right. My only disagreement with Dennett will be that, whereas he regards an autobiography as fictive and consequently false in characterizing its protagonist, I regard it as both fictive and true. We invent ourselves, I shall argue, but we really are the characters whom we invent. (Velleman 2006: 205-206)

This view about the importance of the past for human lives is comprehensive in character, for it relates to a metaphysical debate regarding the reality of the self as a narrator. As such, this view is foreign to a political conception of justice.

But some may raise an objection to my critique of Rawls's forward-looking conception of the person, in § 1.1, by making an analogy between Velleman and Dennett's debate and the one between Parfit and Korsgaard to which I briefly alluded to there. For this latter debate is about the

repercussions on morality of putative metaphysical conclusions about the nature of personal identity. Thus, it may be argued that just as Velleman's claims are improper for a political conception of justice, so are Korsgaard's remarks on the *diachronic unity of agency*. This objection would state, then, that I partly relied on a comprehensive view of agency in order to argue against Rawls's forward-looking conception of the person.

However, the equivalence between Velleman's claims about the reality of the self as a narrator and Korsgaard's remarks on the diachronic unity of agency, at least when it comes to how metaphysically controversial they are, is unfounded, for Korsgaard is, precisely, trying to neutralise the effect of metaphysical disputes in moral discussions. Parfit famously proposes what he considers to be the correct metaphysical understanding of the person, in which personal identity does not hold the importance that is usually ascribed to it. In fact, Parfit claims that personal identity is not a further fact about a person, distinct from and not reducible to 'a brain and a body, and a series of physical and mental events' (Parfit 1984: 223). He then argues that from such a reductionist understanding of a person several things follow regarding morality and rationality, some of them favouring a consequentialist approach to moral assessment (see Parfit 1984: 300-321).

To this argumentative strategy, Korsgaard opposes the divide between *theoretical* and *practical* standpoints. She then endorses the independence of the practical standpoint from its theoretical counterpart:

[W]e must view ourselves in these ways when we occupy the standpoint of practical reason —that is, when we are deciding what to do. This follows from

the mere fact that we must regard ourselves as the causes —the first causes— of the things that we will. And this fundamental attitude is forced upon us by the necessity of making choices, regardless of the theoretical or metaphysical facts. (Korsgaard 1996a: 378)

This way of proceeding is quite close to the spirit of Rawls's political liberalism itself. Moreover, the approach seems to be a development of Rawls's (*CP*: 286-303) explicit defence of the independence of morality from putative metaphysical conclusions. So, against the objection we are considering now, we must notice that it is from the practical point of view that the diachronic unity of agency is asserted. As Korsgaard holds, 'this does not require that your agency be located in a separately existing entity or involve a deep metaphysical fact. Instead, it is a practical necessity imposed upon you by the nature of the deliberative standpoint' (Korsgaard 1996a: 370).

Considering the relation between the citizens and the State, this can be framed in the following way: regardless of whether it is a fact of the world that we are free diachronic unified agents, the State (just as we ourselves) cannot but assume that we are. This claim does not deny or assert that we *in fact* are such type of agents. Thus, here we find a substantial difference with Velleman's claims on the self as a narrator. For him, as we just saw, such a self exists in the fabric of the world and is just as real as any other object in such a world —like my brain. The self as a narrator is thus not (or not only) a postulate of the practical standpoint but, rather, an alleged object of the world described by the theoretical standpoint. Considering this, I think that the analogy between Velleman's self as a narrator and the thesis that agency

presupposes diachronic unity is unfounded. While one is a comprehensive metaphysical claim improper to a political conception of justice, the other is not.

Finally, there is a third concern of comprehensiveness. This has to do with whether the State, by promoting what I have called a critical historical narrative (i.e., a particular understanding of the past that is sensitive to the existence of relevant historical injustices), is already promoting a comprehensive view over others.

The least that can be said against this concern is to point out that, as we saw in § 1.3, every single liberal society creates a historical narrative for itself. Such a narrative is not the result of chance. Several putative facts about the origin and development of society are chosen with the purpose of creating a strong sense of belonging, solidarity, and unity. Acknowledging relevant historical injustices is not different from this non-stopping exercise of creating a historical narrative. Therefore, this third worry, if maintained, would be stating something stronger: creating a historical narrative of *any* kind is always a comprehensive exercise from which a liberal state must refrain.

I called this position in § 1.3 *the libertarian view of history* —since it advocates for exclusively leaving to citizens the creation of historical narratives within a given society. As I said there, I find such a view implausible on several accounts. However, the third worry of comprehensiveness could persist even if I am right about the unlikelihood

and undesirability of a liberal State guided by *the libertarian view of history*. It can be said that the very fact that a liberal State cannot do without promoting a specific view of its history only shows that every single liberal State favours a comprehensive view of history. In this way, this third worry becomes an objection to the possibility of a political conception of justice all together. Yet I think this further claim is also incorrect. This claim puts the competing views about the past on a par with the competing views about the good life. And, indeed, some theorists working on *public memorialisation* sometimes talk in a way that suggests this analogy. For instance, with regards to the dictatorial past of South America in the second part of the twentieth century, Elizabeth Jelin says:

Despite the relevance and centrality of these confrontations on the content of democracy itself, the dictatorial past of the 1970s and 1980s is still very much part of the current debate. Many victims and their advocates demand a complete account of these abuses that took place under dictatorship [...] Others, claiming that they are concerned above all with the functioning of democratic institutions, emphasize the need to focus on the future rather than the past [...] Still others look at the past in order to glorify the 'order and progress' that dictatorships presumably secured. Thus, there are competing and conflicting understandings and memories of the past in societies that are emerging from periods of political violence and state repression. [...] In all cases, as time passes and it becomes possible to establish or conceive a temporal distance between past and present, alternative and even rival interpretations of the recent past and its memories take the center stage of cultural and political debate. (Jelin 2007: 139-140)

This way of framing the struggles taking place in public memorialisation states that there are many competing views about the past. The third worry of comprehensiveness we are exploring exploits this and holds that the State will always favour one to the detriment of the others and in so doing will adopt a comprehensive view about its history.

Against this worry, we need to remember that not *all* conceptions of the good are permissible in a liberal state. Non-liberal conceptions of the good are forbidden and discouraged due to their rejection of the very principles of liberal justice. Rawls is unequivocal about this when he denies that what it may be called absolute neutrality is possible:

No society can include within itself all forms of life [...] As Berlin has long maintained (it is one of its fundamental themes), there is no social world without a loss [...] But these social necessities are not to be taken for arbitrary bias or injustice. (*PL*: 197)

Thus, even if we preserve the analogy between conceptions of the good and historical narratives, we can say that *certain* historical narratives are to be discouraged in a liberal society —just in the very same way in which non-liberal conceptions of the good life are discouraged in a liberal State. In fact, I believe that the moral grounds for the exclusion of certain historical narratives and certain conceptions of the good are exactly the same: the denial of the equal status of all citizens. A historical narrative that denies or eschews the existence of historical injustices simply is a *non-liberal* historical narrative.

For the stated reasons, appealing to the *non-comprehensive character* of Rawls's theory does not justify the lack of engagement with historical injustice. None of the three worries of comprehensiveness (the independence of the personal domain, the avoidance of controversial metaphysical claims, and the role of the State in discouraging certain historical narratives) affect the State's exercise of acknowledging and rectifying historical injustices.

2.4 Concluding Remarks

This concludes my assessment of the view that, despite the importance of historical rectification in the real world, historical injustice is nevertheless beyond the scope of Rawls's theory. As we have seen, this view is mistaken. The *ideal, forward-looking, and non-comprehensive* character of Rawls's theory does not justify its lack of consideration of historical injustice. Discussion amongst Rawlsians should not gravitate around the question of which aspect of Rawls's theory justifies excluding concern for past injustice. Rather, discussion should centre on the question of whether Rawls's theory as it stands right now contains the normative apparatus to respond to them or not. This is the question that I will explore in the following chapter.

3

Towards a Principle of Historical Rectification for Ideal Theory

In the previous chapter we saw why the current default position holding that historical injustice is beyond the scope of Rawls's theory is mistaken. The *ideal, forward-looking, and non-comprehensive* character of such a theory does not warrant the absence of concern for past injustice in Rawls's description of a just society. Therefore, in this chapter I will advance in a more tentative fashion one way in which Rawls's theory could respond to the challenge posed by historical injustice. To that effect, in § 3.1 I first consider how the original principles of justice-as-fairness might relate to the rectification of historical injustice. In § 3.2 I consider how the inclusion of a specific *backward-looking* principle to justice-as-fairness could help to clarify the commitment of Rawls's theory to the idea that *no liberal society is fully just until historical rectification has taken place*. Finally, in § 3.3 I will elaborate on the putative conflict between *retribution* and *recognition*-driven measures within recent historical rectification.

3.1 The Two Principles and Historical Rectification

Every liberal society counts historical injustices within its past. Considering this, every time a new government comes into power, numerous issues regarding the rectification of the past arise. The very first one is, precisely, whether to engage in historical rectification or not —Elster 2006: 6 calls this the ‘fundamental decision’. That this goal can be postponed indefinitely does not mean that the new government is not concerned with meeting liberal-democratic standards. On the contrary, it could be argued that rectification should be partly postponed for that very reason. Urgent tasks such as ensuring the rule of law, strengthening the democratic political system, and diminishing inequalities between different social sectors are all liberal goals of the first order that might be prioritised. A society can even make incredible improvements on all these scores *without* engaging in the rectification of past injustices at all. If this pattern continues for a sufficient period of time, we can imagine a society in which *the fundamental precepts of liberal justice are almost fully met in the **present** even when no rectification of past injustice has ever taken place*. These reflections prompt the following hypothetical scenario:

Scenario 1: Imagine two countries respectively called Argentina and Uruguay whose societies become, after long transitional processes, well-ordered in most respects according to Rawls’s description (see *PL*: 35). First, all citizens accept —and know that the rest accept— the principles of justices-as-fairness. Second, citizens have developed a sense of justice that allows them to comply with the directives of their social institutions. And, third, no *present* injustice occurs within both societies. That is, the institutions of their basic structures are almost perfectly regulated by the principles of justice-as-fairness: there is a

social-minimum ensured for every citizen; all political and civil liberties are equally distributed amongst all citizens; there is fair equality of opportunity in the competition for social positions of advantage in most respects; and, finally, the fact that certain citizens obtain the best social positions benefits the worse-off of society. Now suppose that some victims or descendants of victims in both societies raise the issue of the injustices committed in the late 1970s and early 1980s. Argentina establishes a truth commission, allocates some of its budget or calls for donations to build memorials and museums in honour of past victims, compensates the relatives of citizens abducted and assassinated during the so called *dirty war*, and even rescinds self-amnesties that former aggressors gave to themselves before they left office. By contrast, Uruguay simply decides to draw a line under the past: “let bygones be bygones” is the spirit of its policies towards the past.²²

Under the suppositions of **scenario 1**, the question is: which of the two courses of action is just? Considering the arguments offered in § 1.1-1.3, I hope you would agree that the only just course of action would be the one adopted by Argentina.

²² **Scenario 1** is inspired by the actual and divergent paths that the current Argentinean and Uruguayan societies are taking regarding the injustices committed in the 1970s and 1980s to persons considered political dissidents or ‘subversive’ elements. Argentina has come back to the path it started in 1983 when it established the second truth commission of the contemporary era. It has taken several *recognition*-driven measures such as converting into a memorial the Navy Mechanical School (the biggest torturing centre during the *Dictadura* — 5,000 out of the calculated 30,000 abducted citizens were tortured and killed there) as well as marking the day of the military coup in 1976 as a national day of mourning (March 24). Likewise, Argentina has set in place strong *compensation*-driven measures, passing legislation that allows direct monetary remuneration to victims’ relatives and also to citizens born while their parents were in illegal captivity (see Guembe 2006: 25-47). Finally, Argentina has also undertaken strong *retribution*-driven measures, judging and imprisoning before March 2011 more than two hundred officials involved in human rights violations, including the two presidents of Military Junta, Jorge Videla (1976-1981) and Reinaldo Bignone (1982-1983) (see ‘Argentina Marks Coup Anniversary Amid Dirty War Trail’, <http://www.bbc.co.uk/news/world-latin-america-12832677>). By contrast, Uruguay had refrained until very recently from undoing an amnesty law first approved in 1986. The law was ratified by referendums both in 1989 and 2009 as less than 50% of voters voted against it (see Gillespie 1991: 218-222). In March 2011, the Parliament also failed to overturn the same law because of the decisive votes of some key figures of the *Frente Amplio*, the party in office since the arrival of José Mujica to the presidency. It was not until October 27, 2011, that the stated law was finally overturned by the Parliament, a decision that awaits Mujica’s ratification (See ‘Uruguay overturns amnesty for military-era crimes’, <http://www.bbc.co.uk/news/world-latin-america-15473619>).

Now, it could be argued that although historical injustice is within the scope of Rawls's theory, the effective fulfilment of the principles of justice-as-fairness suffices to ensure its rectification. Accordingly, for this view to succeed, it must be the case that a society in which relevant historical injustices remain un-rectified does not fully realise Rawls's original principles. An obvious first candidate in this respect is the principle of fair equality of opportunity.²³ For, as we have seen, historical injustices may have tremendous effects on the capacity of victims to pursue their plan of life; these effects can introduce unfairness in the competition for social positions of advantage, affecting victims' prospects but not those of the rest of the citizenry.

Coming back to **scenario 1**, the proponents of this view must argue that the principle of fair equality of opportunity is not met in our imagined Uruguayan society. For as long as their past political harms have not been rectified, victims have to compete for social positions at a psychological disadvantage in relation to the rest of the citizenry. In fact, several arguments regarding the connection of historical injustice and fair equality of opportunity have been made in the literature on positive discrimination. Paradigmatically, Sher holds that the very moral basis of affirmative action is

²³ Rawls's final formulation of this principle, in conjunction with the difference principle, reads: 'Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and, second, they are to be to the greatest benefit of the least-advantaged members of society' (*JF*: 42-43).

the fact that such policies restore equal access to positions of advantage amongst victims of past injustice. Sher states:

It is, I think, the key to an adequate justification of reverse discrimination to see that practice, not as redressing of *past* privations, but, rather as a way of neutralizing the *present* competitive disadvantage *caused* by those past privations and thus a way of restoring equal access to the goods which society distributes competitively. (Sher 1975: 163. Italics in the original)

There are three important things to notice about this claim. First, that affirmative action is understood in a ‘forward-looking’ way (see Sher 1999: 85-86 and Taylor 2009: 478). As Sher holds in the passage quoted, this approach is not seen as a way to rectify historical injustice as such, but only its *effects* on citizens’ present opportunities. Accordingly, by exclusively trying to ensure equality of opportunity in the present, affirmative action policies are not thought to accomplish full rectification for past wrongs (see Sher 1977). Second, this account of affirmative action is not based on the idea that such policies must be in place because they guarantee a more diverse society (see Sher 1999). Third, affirmative action policies are understood as *temporary* measures to achieve the stated objective (fair equality of opportunity amongst victims and the rest of the citizenry). In this sense, we can think of such policies as one of the strategies to arrive from our disordered societies—in which fair competition for positions is not guaranteed due to the unfair historical background of some citizens—to well-ordered societies—in which the citizens’ background does not influence competition. Thus, once the transition to a well-ordered society is completed,

there is no need for maintaining affirmative action policies. This is also the spirit of Samuel Freeman's following remark:

So-called "affirmative action", or giving preferential treatment for socially disadvantaged minorities, is not part of FEO [fair equality of opportunity] for Rawls, and is perhaps incompatible with it. This does not mean that Rawls never regarded preferential treatment in hiring and education as appropriate. In lectures he indicated that it may be a proper corrective for remedying the present effects of past discrimination. But this assumes it is temporary. Under the ideal conditions of a "well-ordered society", Rawls did not regard preferential treatment as compatible with fair equality of opportunity. (Freeman 2007a: 90-91)

The work of these authors supports the main claim of the view we are considering: relevant past wrongs have present effects on the opportunity of victims, and because of that a society that has not rectified such effects is not well-ordered. Thus, even when no injustice is committed in the present, the effects of past injustices on the opportunities of victims are still present in our imagined Uruguay. Indeed, the preceding point must be fully granted.

However, it would be a mistake to think that the normative significance of historical rectification is exclusively related to the fact that society should guarantee equal opportunity to their citizens. To see this, consider **scenario 2**:

Scenario 2 is just as **scenario 1** but with one difference. Most victims of past injustice are highly resilient psychologically speaking. In fact, a former political dissident, call him José Mujica, even becomes president of the well-ordered Uruguay. During the 1970s and 1980s, Mujica was seized and tortured by the repressive forces of the dictatorial regime for up to fourteen years. By contrast, after the consolidation of democracy in the Uruguayan transition and establishment of a well-ordered society, he was obviously neither imprisoned nor tortured; and —as the kind of *Übermensch* that he is— both his ability and his willingness to compete for public office remained intact. He himself states this very clearly. Suppose, however, that he does say something along the following lines: 'You see, after democracy was re-established, I really had no disadvantage in my

struggle to eventually become President. My society offered all opportunities and I took them as they were coming. I did not resent my past while in such a process. However, I still think there is something wrong with what was done to me during the *dirty war*, and I believe that my society should officially acknowledge that. I think it is the prerogative of persons like me to stake this claim to recognition'.²⁴

The point to underscore in **scenario 2** is that Mujica can raise the complaint that, to the extent that his past remains un-rectified, even when such a past did not affect his opportunities ever since the Uruguayan society became well-ordered (after all, he even managed to become president)²⁵, an injustice is still taking place. Such injustice is related to two other normatively salient consequences of a non-rectified historical wrong.

First, as I argued in § 1.2, when the State fails to provide means for rectification of past violence, it fails to show respect for victims' moral equality. **Scenario 2** illustrates that the State should provide these means simply because of that fact. The priority of this goal over fair opportunity is shown by the fact that even when victims' opportunities are not affected the State *still* must offer means for rectification. In this sense, in **scenario 2** Mujica is demanding what Borneman 2005: 60 calls a 'retroactive recognition of dignity' —beyond the equal opportunities he must enjoy.

²⁴ **Scenario 2** is inspired by the current Uruguayan society, with the difference that such a society is not well-ordered and president Mujica oscillates on whether to pursue institutional rectification for the gross injustices committed against him and other citizens like him or not.

²⁵ This is not to say that, more generally, the fact that a person from a historically disadvantaged group manages to attain a very high social position in a given society proves that fair equality of opportunity has been provided in such society. For a person from a historically disadvantaged group can indeed attain such a position despite the fact of her disadvantage with regard to other citizens.

Of course, this seems a rather unusual case. For **scenario 2** exploits the psychological strength of Mujica and the imagined Uruguayans, strength that is by no means common. But notice that Mujica can still say something similar *even if his non-rectified past affected his opportunities*. He could say that rectification is *not only* due to the fact that his opportunities in becoming president were affected. These opportunities represent just one aspect of his moral standing as a person that is supposed to be protected by the prerogatives of justice. Another aspect of his moral standing as a fully equal person, having to do with the dignity of his life as a whole, is also to be protected by such prerogatives. The official recognition of his terrible past mistreatment is thus thought to restore this other aspect of his moral standing.

Second, Mujica's compliant in **scenario 2** is also related to the fact that —beyond the importance of providing equal opportunity— rectifying a historical injustice is central for ensuring the liberal integrity of society as a whole. Since Mujica's past mistreatment has not been acknowledged as something *wrong* that *should not have occurred*, it is not clear that the Uruguayan society shows *regret* for the abuses committed during its non-liberal period. Some might even hold that such a period was necessary for Uruguay to be where it is in its present. In fact, as we will see in § 3.3, this is precisely what many militia members who took part in gross human rights violations during Uruguay's *dirty war* actually believe. By failing to acknowledge that Mujica's mistreatment was *wrong* and *should not have*

occurred, it is not clear then that citizens have distanced themselves from the unfortunate belief that illegal detentions, torture, and forced disappearances *might* be a necessary means of accelerating the process of achieving a well-ordered society.

These reasons show that the aim of providing means for historical rectification is more fundamental than the need to secure a fair competition for positions of advantage. Considering this, it could be argued that it is not the perfect regulation of the fair equality of opportunity principle that is missing in **scenario 2** but, rather, the perfect regulation of Rawls's first principle of justice.²⁶ In particular, it could be defended that the lack of official acknowledgement of Mujica's past mistreatment deprives him from what Rawls (*TJ*: 53) calls the, 'freedom of the person, which includes freedom from psychological oppression' (see also *PL*, 291, *JF*: 44), and because of that, such a liberty has been distributed unequally in our imagined Uruguay.

This last claim is more plausible. Yet I am inclined to think that this theoretical way of stating that means for historical rectification will be provided in **scenario 2** places too much explanatory weight on the importance of guaranteeing the equal distribution of the freedom from psychological oppression. Furthermore, the importance of historical rectification in **scenario 2** goes beyond the fact that Mujica has been *individually* psychologically oppressed —conceding that this is the case, for in

²⁶ Rawls's final formulation of this principle reads: 'Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all' (*JF*: 42).

scenario 2 Mujica is presumed to have tremendous psychological strength. As we saw, the importance of rectification in this case is also related to the fact that the Uruguayan society must show institutional regret for its terrible non-liberal period of abuses.

That this route for ensuring full historical rectification is not entirely satisfactory might be clearer by considering the following scenario:

Scenario 3: Suppose that Mexico, after a long transitional process, becomes a well-ordered society. Basic institutions are fully regulated by Rawls's principles of justice-as-fairness and all citizens know, accept, and are motivated by this fact. Now suppose that some citizen activists raise the complaint that the State must acknowledge that during the 19th century the Mexican liberal State tolerated the extreme exploitation of thousands of indigenous persons at the beginning of its long transition to becoming a well-ordered society. There is no direct relation between the activist and those indigenous persons (the complainants may not even be indigenous persons themselves) and yet the former think there is something wrong in not recognising such a past. Notice that, since we are supposing that Mexico is a well-ordered society in this scenario, all indigenous persons in the present, as citizens of the Mexican State, are now effectively protected by the prerogatives of justice-as-fairness, including the fact that they can compete in fair equality of opportunity for social positions of advantage. Yet, the complainants ask the State to advocate for what in § 1.3 I called a critical historical narrative. That is, a narrative that not only celebrates the putative remarkable facts about the origins of the Mexican State in its transition to becoming a just society but, also, one that is sensitive to the past wrongdoing of Mexican society as a cultural and political continuum over time.²⁷

²⁷ **Scenario 3** is inspired by the current Mexican state of affairs. The gross exploitation of the indigenous population throughout Mexico's history includes specific acts of genocide, systematic confiscation of land and property, systematic exclusion from political participation and discrimination (see Robins 2005: 23-68, Vandervort 2006: 138-161, Gueda 2010: 265-285). Against this well-documented reality, time and again organised groups of citizens ask not only for providing equal participation to indigenous persons in current political proceedings, but also for the official promotion of a multicultural understanding of the nation, including the claim that history lessons in primary schools should acknowledge the systematic exploitation of indigenous communities throughout Mexican history. This demand, amongst many others, has been set forward more directly by militants of EZLN

As we saw in § 1.3, historical rectification of distant past wrongs within a liberal society requires the adoption of a critical historical narrative. This is precisely what citizens in **scenario 3** are demanding. However, it seems hard to defend in this case that the very citizens complaining about the lack of rectification are been psychologically oppressed. These citizens are supposed to have successfully achieved their plan of life in a context in which most prerogatives of justice are fulfilled *except* for the absence of the stated act of historical rectification; we have even assumed that they are not indigenous persons themselves. Arguing that the absence of rectification in this case is an infringement on citizens' freedom from psychological oppression seems to stretch too much the way in which such a freedom can be violated.

Even when a non-critical historical narrative might not directly oppress citizens taken individually, there are still important reasons for rejecting it. As we saw in § 1.3, a liberal society is a cultural and political continuum that uses institutional means to promote a particular understanding of its history. Institutions must then express their liberal commitment by stating their condemnation of the non-liberal institutional context that allowed the perpetration of political violence in the past. For a society that systematically fails to acknowledge the past wrongdoing that occurred within its long transition to become a well-ordered society undermines its own liberal integrity. There is room for some disagreement

(Zapatista National Liberation Army), a political movement originating in the Mexican State of Chiapas in 1994 and now with followers all over the country.

regarding some particular events and their significance, especially when such events become more distant in time. However, we should note that the systematic omission of past injustice promoted by a non-critical historical narrative goes far beyond disputes about what happened at a given time.

Appealing to the first principle cannot directly support the adoption of a critical historical narrative in **scenario 3**. On the other hand, as we saw in **scenario 2**, the entire burden of justification for securing means for recent-past rectification might fall under Rawls's freedom from psychological oppression —as **scenario 2** is meant to show that a society can make incredible achievements in ensuring the rest of Rawls's basic liberties, fair equality of opportunity and beneficial inequalities *without* engaging in the rectification of the recent-past at all. Of course, this is nothing more than a preliminary assessment of how the original principles of justice-as-fairness might directly command the rectification of past wrongs. However, I hope it serves at least to motivate the need to consider a specific principle of historical rectification as a way of tackling the normative significance of historical injustice within justice-as-fairness.²⁸

3.2 Towards a Principle of Historical Rectification for Ideal Theory

As **scenarios 2** and **3** try to exemplify, a society can make incredible improvements in important liberal-democratic scores *without* engaging

²⁸ In § 4.2 I will come back to the relation between the original principles of justice-as-fairness and historical rectification in the context of a comparison with Seana Shiffrin's claim that such a conception of justice should be complemented by a specific anti-discrimination principle.

directly with the rectification of past injustices. This possibility gives rise to the following theoretical positions:

- (i) Ideal liberal justice requires ensuring the prerogatives of justice from the present on.
- (ii) Ideal liberal justice requires ensuring the prerogatives of justice from the present on *and* providing means for historical rectification —such as *retribution*, *compensation*, and *recognition*-driven measures on behalf of victims of past political violence, as well as the institutional promotion of a critical historical narrative.
- (iii) Ideal liberal justice requires ensuring the prerogatives of justice from the present on. It does not require providing means for historical rectification.

I have defended (ii) by pointing to the normative significance of coming to terms with a past shaped by political violence regarding (1) citizens' psychology and self-respect, (2) their mutual respect, and (3) society's liberal integrity. One way of expressing commitment to (ii) is by holding that the basic institutions of a well-ordered liberal society are also regulated by a specific backward-looking principle ensuring that processes of historical rectification take place. This makes it clear that, from the point of view of justice, historical rectification is not optional even when most liberal-egalitarian goals are fulfilled.

As I also tried to show in the previous section, while it is clear that Rawls's description of a well-ordered society is an expression of (i), it is not so clear at the present moment how such a description is a robust expression of (ii). Let me then resume in four related points the case in favour of including a specific principle of historical rectification to justice-as-fairness.

First, the three considerations on which I based the normative significance of historical rectification —(1) citizens' psychology and self-respect, (2) citizens' mutual respect, and (3) society's liberal integrity— are of incredible importance for Rawls's theory. Second, as I argued through § 2.1-2.3, the theoretical constraints on which such a theory is built do not exclude concern for historical injustice. Third, the inclusion of this principle could help to create theoretical balance within justice-as-fairness, for the importance of fulfilling the prerogatives of justice in relation to the past, present, and future of a well-ordered society would be fully addressed (considering that such a society is a cultural and social continuum with an intergenerational span). And, fourth, including such a principle within justice-as-fairness is *not* a violation of the Rawlsian divide between ideal and non-ideal theory. This principle would simply clarify the commitment of Rawls's theory to (ii) — that is, commitment to the idea that *no liberal society is fully just until historical rectification has taken place*. In fact, Simmons's remarks reviewed in § 2.2 offer strong support for this fourth point:

Rawls's ideal theory, it is said, pays no attention to the long histories of injustice [...] But the kind of "oversights" with which such criticism charge Rawls are not in any way essential to Rawls's characterisation of ideal theory or its relationship to non-ideal theory; they are rather simply a function of the specific approach that Rawls employs to derive the content of ideal theory, to argue for his particular, favored conception of "justice as fairness". Historical injustice goes unaddressed in Rawls's theory because the derived principles of justice are purely "forward-looking", because the choice problem given to Rawls's original position contractors requires their choice of forward-looking principles. (Simmons 2010: 32-33)

I have tried to show that Simmons is mistaken in arguing that the choice problem given to the parties in the original position justifies the selection of

exclusive forward-looking principles. However, a few lines below the passage previously quoted, Simmons states the following idea:

A quite different content to ideal theory (for instance, one that was more sensitive to the need to redress historical injustice) could be defended while still subscribing wholeheartedly to the Rawlsian version of the ideal-nonideal distinction. (Simmons 2010: 32-33)

Simmons is right in arguing that a liberal theory of justice *can* show concern for historical injustice *without* thereby violating Rawls's distinction between ideal and non-ideal theory. Such a theory would subscribe to (ii) as the correct ideal of liberal justice at which we should aim in our actual social proceedings. Once such a commitment is established, different specific approaches to the *transition* to the fulfilment of that ideal can be provided. Of course, we should expect countless variations of appropriate rectificatory measures depending on the context of the society in question. How many persons were the target of wrongdoing, for what arbitrary reason, how grave the wrong inflicted on them was, and how distant in time the stated injustice is, are all contextual elements that propel different institutional responses for each case. A theory of justice on a philosophical level of generality cannot but come short of offering more concrete stipulations in this regard (see Elster 2004: 78 and de Grief 2006b: 466).²⁹

It is nevertheless worth mentioning that specific rectificatory frameworks of recent-past political violence usually require of a coherent set of policies including instances of each of the three measures outlined in (ii).

²⁹ Elster states: 'I do not aim at presenting a "theory of transitional justice". As in my earlier work on local justice, I have found that context-dependence to be an insuperable obstacle to generalizations' (Elster 2004: 78).

For instance, when only *retribution*-driven measures are undertaken, victims might think that the State is merely focusing on a struggle *against* aggressors and is making no effort on *victims' behalf* directly (see de Grief 2006a: 2). Likewise, if the State focuses exclusively on economic compensation, victims might perceive this as a form of 'blood money', that is, as a way of silencing their claims for justice in exchange of economic advantage (see Elster 2004: 166n). If the State implements only *recognition*-driven measures, victims might perceive this as a *facade* for covering aggressors' lack of real accountability for their past wrongdoing (see Verdeja 2006: 130-31). Accordingly, regardless of the contextual variations expected in different liberal societies, effective programs of recent-past rectification must include a coherent set of *retribution*, *compensation*, and *recognition*-driven measures. When one of these measures is absent, this is in need of justification.

Also, it is important to note that, in general, the more recent the historical injustice is, the more relevant it becomes and the more negative effects it has on the society in question (see Sher 1981: 6 and Elster 2004: 222-229). Considering this, a principle of historical rectification within ideal theory could state that while recent cases of historical injustice call for strong *retribution*, *compensation* and *recognition*-driven measures as appropriate, other cases involving distant past wrongs call for *recognition* within the critical historical narrative promoted by the institutional means of the State (see Thompson 2001: 132-135). For *retribution*-driven measures rectifying the course of history are simply impossible. Likewise, *compensation*-driven

measures going all the way to the accepted historical origins of a liberal society make little sense —if any sense at all. However, as we have seen, this does not mean that history as such is out of the scope of liberal institutions' moral assessment.

In sum, the main reason for including a specific backward-looking principle within a liberal ideal theory of justice is the capital importance of stating that (ii) represents the ideal we should aim for in our actual social proceedings. In Rawlsian fashion, this idea can be formulated by saying that a liberal well-ordered society is also regulated by a principle of historical rectification. Accordingly, what Elster calls the 'most fundamental decision' should be always resolved in the same way:

The dependent variables [of historical rectification] may be conceptualized as a series of *decisions*. The most fundamental is the decision whether to address the wrongdoings of the past at all, or rather draw "a thick line" through the past. If the former option is chosen, the new regime may weight the options of *truth* or *justice*. (Elster 2006: 6. Italics added)

Elster makes this remark in the context of rectification within post-conflict societies. However, as I said at the beginning of § 3.1, every new government that comes into power in all liberal societies confronts such a fundamental decision —since every society counts historical injustices within its recent and distant past. From the point of view of justice, drawing 'a thick line' through the past or 'closing the books' is not an option for any liberal society. In the following section I would like to address the second putative decision confronted by a society that decides to set forth mechanisms of historical

rectification. This decision is, nevertheless, more pressing for societies in need of rectifying cases of severe recent political violence.

3.3 On a Recurrent Transitional (non-Ideal) Problem: *Retribution VS Recognition*

Admittedly, as Elster holds, some rectificatory measures might come into conflict when dealing with cases of recent political violence. Consider what scholar Marie Breen Smyth, quoted in § 1.2, states:

Whilst the provisions of amnesty may be a valuable incentive to perpetrators to participate in a truth recovery process, the provision of amnesty precludes the achievement of justice for victims [...] Victims in general are seemingly faced with a choice between more information or 'truth' on the one hand and justice on the other. (Smyth 2007: 13)

In this section I would like to address this conflict in some more depth. I focus on this conflict, rather than on other problems that theories of rectification confront, for two reasons. First, this conflict is quite prominent in the literature (see, for instance, Van Zyl 1999, Elster 2004: 118, Lutz 2007: 326, Smyth 2007: 6-21, Roper and Barria 2009). Second, I think this conflict is indeed about what is required for *doing justice* to victims. This latter feature distinguishes the clash between different rectificatory measures from the other frequent and overwhelming problems of rectification theories. For most of such problems have to do with *how* to reconcile the requirements of justice with the current practical possibilities of rectification in particular societies.

The clash of *retribution* and *recognition*-driven measures was fully assumed by the paradigmatic truth commission: South Africa's Truth and

Reconciliation Commission (TRC). For the TRC offered 'amnesty' to aggressors who were willing to fully and publicly disclose the nature of their wrongdoing. The motivation behind these amnesties, which rub against the very idea of retributive justice, was to evince the confessions of aggressors. Along with the testimonies of the victims, these confessions would be part of the material used to construct a critical historical narrative about the past (see Holtermann 2010). In similar fashion, the purpose of the Guatemalan Truth Commission, tellingly called the Commission for Historical Clarification (CEH), was not to impose sanctions against the culprits involved, and therefore it refrained from 'individualizing' responsibilities (see Quinn & Freeman 2003, Chapman & Ball 2001). Instead, the CEH was expected to write a report about the legacy of past violence, which among others things characterised the State-commanded decimation of Mayan people during the civil war as *genocide*; in the words of one of its commissioners, the report contained pages covering the darkest period in the history of the country, as well as their causes, which could be traced back as far as the 'Spanish conquest' and 'the poisoning effect of racism on Guatemalan society' stemming from it (see Tomuschat 2001).

But the importance of building a critical historical narrative is not the only normative reason for focusing on *recognition* rather than *retribution* when rectifying a recent past injustice. In some cases the individuation of moral responsibility required for applying correct *retribution*-driven measures might be extremely difficult. An important distinction in this

regard is frequently missed in the literature of transitional justice: that of the different levels of responsibility of *high* and *low*-rank aggressors.

There is one striking feature of numerous cases of massive political violence: aggressors' belief that they are doing the right thing by persecuting and depriving victims. Most militia members of all dictatorships in Latin America during the period of 1960-1990 who have been forced to accept their leading participation in gross human rights violations, reiterate this belief (see Payne 2008: 141-172). Consider, for instance, what colonel Gilberto Vázquez (prominent member of the Uruguayan militia who abducted, tortured, and disappeared political dissidents —now in jail in Argentina) stated in a recent interview to the local newspaper *Ultimas Noticias*:

'Torturing is horrible and stays in your consciousness for the rest of your life. However, it was necessary [...] It was as when a surgeon has to amputate a leg because with gangrene the person dies. He does not do so because he likes it [...] So I do not regret doing it. I am proud of my participation in the salvation of my country'. (Reproduced in *La Jornada* 1/02/2011. Translation mine.)

The immediate task then becomes to explain how aggressors can form such an attitude. We need to avoid ingenuous and meaningless explanations such as their putative inner-immorality or evilness. This is, I believe, perhaps the most valuable lesson from Hanna Arendt's analysis of Eichmann's trial (see Arendt 2006). Rather, a more accurate explanation of the stated belief has to do with noting a previous process of normative distortion affecting aggressors' moral judgement and subsequent actions.

There are two important elements that inform moral judgement: the processing and pondering of information that calls for moral evaluation and moral concepts themselves. Barbara Herman (see Herman 1993, 2007) calls the conjunction of these two elements the ‘rules of moral salience’. These rules determine which situations are in need of normative assessment (and which factual information is to be considered) as well as how we understand moral concepts such as justice, equality, reciprocity, etc., *before* we actually exercise our moral judgement. For Herman, the correct realisation of such a capacity requires the implementation of the categorical imperative in the form of the ‘CI procedure’, that is, a procedure testing the permissibility of our actions by envisioning the universalization of the maxims that prompt them (see Herman 1993: 75). Regardless of whether this characterisation of the exercise of moral judgement is correct or not, what I want to underscore is that Herman’s analysis makes it explicit that such a capacity is exercised *within a social context*. The rules of moral salience within which agents make moral judgements are passed on ‘as elements in a moral education’ and, more generally, ‘as part of socialization’ (Herman 1993, 77-78). It is a further advantage of Herman’s analysis that it identifies this context as fashioned by *rules*, for this makes it clear that there is some normative force that induces agents to comply with it.

With this analysis in hand, we now can see how aggressors typically form the belief that they are doing the right thing by partaking in cases of mass political violence: the social context within their moral community has

distorted the rules of moral salience up to the needed point. Exposure to intense ideology plays a mayor role in the achievement of this goal. In line with the first element shaping the rules of moral salience, empirical information is usually distorted up to the point in which partaking in such atrocities is perceived as necessary to ensure the very survival of the community, institutional order, or country. In line with the second element shaping the rules of moral salience, key moral concepts such as equality and reciprocity can be twisted up to the point in which victims are no longer subjects to equal moral consideration. The result of these antecedent processes of normative distortion is such that aggressors evaluate their actions in line with the precepts of morality (see Pauder-Studer & Velleman 2011). Accordingly, we can say that aggressors' actions may become a result of what Gideon Rosen 2003: 64 calls 'blameless moral ignorance', that is, a case in which the agent, due to social factors delivering incorrect factual and normative information, acts wrongly without bearing full moral responsibility for her action or the ignorance that motivated it.

To these social processes affecting agents' moral judgement and actions we should add some psychological ones too. In particular, the empirical evidence provided by the famous Milgram experiments (see Milgram 1974) and their multiple replicas support the claim that agents have a *tendency to obey authority* when they think such an authority is backed up by a solid institutional framework, even in cases in which the directives received conflict with their previously accepted moral standards. Yet further

‘situational forces’ hold a direct effect on agents’ behaviour. John Sabini and Maury Silver have recently presented an excellent review of social psychology studies detailing such forces (see Sabini & Silver 2005). These studies provide evidence of a *tendency to lose one’s moral compass* in a slippery slope manner when confronted with directives of institutional authority (Sabini & Silver 2005: 549), a *tendency to conform with the social context* even when one’s cognitive worldview is drastically at odds with the social consensus (Sabini & Silver 2005: 554-555), and a *tendency to diffuse responsibility* in a context in which others agents might offer altruistic help or even help for ensuring one’s own safety within a group (Sabini & Silver 2005: 555-557). They summarise these findings as follows:

Social psychology since the Second World War has indeed discovered [...] something important about character and morally significant action. The tradition that begun by Ash [whose work evinces agents’ tendency to *conform with the social context*] and continued through Milgram [whose work evinces agents’ tendency to *obey institutional authority*] and Lactané and Darley [whose work evinces agents’ tendency to *diffuse responsibility*] has revealed just how weak, morally weak, we are when confronted with a resolute authority or a unanimous group of other seemingly normal people who seem to see the social, moral, and even physical world differently from the way we do. This weakness is partly cognitive —people tend to loose their moral compass— but is also partly a matter of people’s being unable or unwilling to expose themselves, to disrupt social situations, by exposing their different perceptions of the world. (Sabini & Silver 2005: 560)

Despite their acknowledgement of the incredible force of these tendencies, Sabini and Silver are careful enough not to accept the claim that such findings show that moral character is simply and globally defeated by the social context. Rather, they argue that certain contexts enhance the strength of such tendencies and that, by being aware of them, agents can mitigate their effect on their own actions (Sabini & Silver 2005: 562).

The pervasive effect of these social and psychological factors in contexts allowing the irruption of political violence (i.e., the fact that the rules of moral salience normatively command such an irruption and agents' psychological tendencies to obey institutional authority, conform with the social context, and diffuse responsibility), make the correct individuation of responsibility for adequate retributive action against some aggressors difficult. However, in line with Sabini and Silver's precautions stating that such findings do not defeat moral character all together, I do not claim that individual ascription of moral responsibility is all together impossible. Because of this, I disagree with Rosen when he states:

[I]t makes little sense to hold this injustice against the perpetrator when it would have taken a miracle of moral vision for him to have seen the moral case for acting differently. *It may be hard to believe that moral evil might turn out to be, in the relevant sense, no one's fault.* But so long as the underlying ignorance is no one's fault, it seems to me that is just what we should think. (Rosen 2003: 66)

Rosen thinks that the inescapable effect of the social context in informing agents' moral judgement and subsequent actions poses an incredible challenge to the ascription of moral responsibility in a great majority of cases (see Rosen 2003, 2004). However, at least when it comes to cases of political violence, it is hard to believe that wrongdoing is 'no-one's fault'. For it is extremely important to note that most of the time the distortion of the rules of moral salience is commanded by the high or influential hierarchy of society. These hierarchies usually have all the necessary means to avoid the stated distortion. They have access to the relevant empirical and normative information showing that massive political violence against a particular

group is wrong. In fact, such groups might even have access to confidential information evincing the opposite of what they command. The fact that high-rank aggressors have access to such information shows that their actions are not a result of 'blameless moral ignorance'. Likewise, we should also notice that high-rank aggressors are not under the pressure of authority, as is the case with low-rank aggressors, and thus they cannot invoke the normal psychological tendencies to follow institutional authority against previously accepted moral standards described by the Milgram experiments. Therefore, while ascription of moral responsibility to low-rank aggressors might be difficult, ascription of moral responsibility to high-rank aggressors is indeed quite possible. Thus retributive action against the latter is usually justified. Moreover, the fact that high-rank aggressors shift the rules of moral salience and command authority in order to allow the irruption of political violence exacerbates their moral responsibility for it, justifying an even more severe retributive action against them.

These remarks are only aimed at showing that it is not always true that *doing justice* to victims of recent political violence requires *retribution*-driven measures against all aggressors. Some cases may require instead the implementation of both *compensation* and *recognition*-driven measures. This conclusion goes against two extended and related views regarding post-conflict violence rectification: that policies granting retributive amnesty to aggressors are *always* in detriment of doing justice to victims, and that only prudential or pragmatic reasons can justify such policies (such as the fact

that retributive action against aggressors at a given time might conflict with stability and reconciliation). Such views are mistaken at that level of generality. For, as we have seen, doing justice to victims might not require *retribution*-driven measures against low-rank aggressors at least in some cases. The reason for this, as we have also seen, is moral rather than prudential in character: the fact that low-rank aggressors' individual moral responsibility for their deeds may be extremely hard to pin down.

This does not mean that low-rank aggressors can fail to accept their deeds as properly theirs. For, once the social context is restored for ensuring the equal integrity of all citizens, low-rank aggressors can see that their past deeds, although perhaps predictable within the previous normatively distorted social context, were fully and plainly wrong. Public apologies, another important *recognition*-driven measure, have the purpose of stating this belief. Amongst other things, a sincere public apology restores the equal status of a victim by acknowledging her past deprivation as a wrong; it is also a statement of the commitment not to repeat the stated wrong and, more generally, a strong way of showing repudiation for the illiberal social context that allowed it (see Verdeja 2010: 570-571). When a public apology for past wrongdoing is not offered, it is simply not clear that the aggressor sees her past deeds as wrong. This is the case of most militia members of all dictatorships in Latin America during the period of 1960-1990, such as the Uruguayan colonel Gilberto Vázquez quoted above, for they refuse to apologise for their terrible deeds. In fact, it is quite common to see high-rank

aggressors defending their past actions and showing no regret whatsoever for committing them (see Payne 2008: 141-172). This means that there persists the belief that victims are indeed not worthy of equal moral consideration. It means, then, that the rejection of a liberal context of moral interaction persists.

3.4 Concluding Remarks

This concludes my assessment of the clash between *retribution* and *recognition*-driven measures, as well as of the theoretical advantages of including a backward-looking principle within justice-as-fairness. As I said, doing so would be a straightforward way of establishing the commitment of Rawls's theory to the ideal that no liberal society will be fully just until historical rectification has taken place. Note, nevertheless, that I am not suggesting that there is no plausible way of arguing in favour of full historical rectification by appealing to the original principles of justice-as-fairness. Yet it seems at least theoretically convenient to include a specific backward-looking principle in order to express direct commitment to that goal. Other authors, however, have made similar claims regarding the silence of Rawls' theory of justice with respect to racial and gender relations. In the following chapter, I will revise some of these claims in the context of the current debate on the correct understanding of equality and egalitarian justice. The purpose of this assessment is to see the different strategies needed to ensure the respect of citizens' moral equality in a well-ordered society. Also, reviewing

these claims help to clarify when it is justified to argue that a further principle should be included within the framework of liberal ideal theory and when is not.

Part II

EQUALITY AND PRACTICE-DEPENDENCE

4

The Protection of Citizens' Moral Equality

In the previous chapters I have criticised Rawls's theory of justice for its lack of concern for historical injustice. In this chapter, I will focus on a similar critique that it is often made against such a theory: its lack of concern for racial and gender relations. In § 4.1 I argue that one way this critique can be framed (which I call the *no-use critique*) fails once we stress the importance of clarifying what I call the fundamental egalitarian aim. In § 4.2 I argue that another way the race and gender critique can be framed (which I call the *institutional injustice critique*) also fails. I focus on Seanna Shiffrin's version of this critique in order to assess her claim that a specific anti-discrimination principle with forward-looking character should be included within justice-as-fairness. Finally, in § 4.3 I argue that a response to a third way the race and gender critique can be framed (which I call the *non-institutional injustice critique*), requires the inclusion of an egalitarian ethos, informing citizens' personal interactions, within the description of a well-ordered society. Importantly, I defend that Rawls's theory already contains all the normative

elements required for establishing that such an ethos is informed by a norm commanding respect for citizens' moral equality.

4.1 The *No-Use Critique* and the Fundamental Egalitarian Aim

Rawls's theory of justice has been attacked for its blindness to several important problems concerning, amongst other things, racial relations (see Shiffrin 2004) and gender relations (see Schwartzman 2006).

These critiques can take numerous forms. The most basic form gives rise to what I will call the *no-use critique*: to wit, the idea that Rawls's theory, by merely offering an *ideal* (in the non-technical sense —see § 2.1) of how a just society should be (i.e. the realist utopia of a well-ordered society), is of no use for tackling the elimination of oppressive social structures of power pending upon women and racially discriminated groups and, what is worse, might even obscure the urgent need in all liberal societies of doing so. Notice that the *no-use critique* is not, or at least not only, directed towards the use of the stipulations of *strict-compliance* and *favourable conditions* when doing political philosophy. Rather, the primary focus of this critique is the fact that Rawls's theory is concerned with offering a description of how an ideally just society should look (see Mills 2005 and Sen 2009 for variations of this critique).

As I have suggested throughout this thesis, I do not find this critique compelling. One of the main tasks of political philosophy (as a branch of practical philosophy more generally) is to offer descriptions of *how things*

should be from a normative point of view. This exercise is not futile, considering that in many cases there is strong disagreement about *how things should be*. For instance, persons (including philosophers, politicians, and citizens) disagree on how much income-inequality is allowed by a just distribution of the burdens and benefits of social cooperation. This disagreement is about how a just distribution *should be*. It is about the *ideal* we should try to realise in our real distributions. Only when we know the *ideal* we want to achieve in this regard (say, for instance, that our distribution should only allow income-inequality in accordance with the difference principle), can we then inquire both *how far* our current distributions are from that ideal and *how* we can adjust them to better approximate the ideal. Taking these two further steps is not possible without knowing what the ideal is that we wish our distributions to fulfil.

There are infinite examples like the previous one in many other domains. To mention another one, consider the current dispute in Uruguay on whether to rectify the historical injustices of the late 1970s and early 1980s. The Uruguayan society is sharply divided in this regard.³⁰ By 2009, nearly half of the politically active citizens expressed that the current state of affairs, in which full rectification of such injustices is forbidden by law, is *not the way things should be*. Conversely, the other half of the politically active

³⁰ As I reported in § 3.1, in 2009 Uruguay held a referendum on whether to nullify the *Ley de Caducidad de Pretensión Punitiva del Estado* (first approved in December 22, 1986), establishing a statute of limitation for crimes committed by military forces and police prior to March 1, 1985. In the stated referendum, 47.98% of participants voted for nullifying the Law and 52.02% voted in favour of preserving it.

citizens expressed that such state of affairs is *the way things should be*. This is a normative dispute at the core of current politics in Uruguay concerning how to deal with its recent past, and one of the fundamental aims of political philosophy must be to clarify what justice requires in this and similar cases. The relevance of defending a principle of historical rectification, as one of the principles regulating the basic structure of a well-ordered society, is precisely that: i.e., stating that no society would be fully just until historical rectification takes place.

Of course, once we agree on the ideal to pursue in each particular case, the pressing issue of *how* to realise that ideal considering the current conditions of the society in question arises. However, as I have said, this further issue only arises once we know what we want to achieve. In this sense, *the no-use critique* subsumes all possible normative work to transitional theory, without realising that any transitional exercise requires an ideal at which we wish to arrive.

The importance of clarifying the ideal we want to arrive at also illuminates the main problem of the so-called luck-egalitarianism, when taken as a fundamental conception of egalitarian justice. Consider what Kymlicka (wrongly attributing to Rawls, as we will see) states in this regard:

...what it would be for a distributive scheme to fulfil the basic aims of Rawlsian theory: a distributive scheme that respects the *moral equality* of persons by compensating for unequal circumstances while holding individuals responsible for their choices. (Kymlicka 2002: 86. Italics added)

Also consider what, in similar fashion, Kok-Chor Tan has stated more recently:

For luck-egalitarians, the idea of the *moral equality* of persons requires that each person take the responsibility for her choices and assume the cost of these choices. Conversely, it holds that no one should be worse off just because of bad luck. (Tan 2008: 665. Italics added)

Notice two important things regarding these passages. First, that at least some luck-egalitarians think that the fundamental normative aim of egalitarian justice is to safeguard citizens' moral equality (see also Dworkin 2000: 3ff for a variation of this statement). In this, I think they concur with so-called political or social-egalitarians (such as Anderson 1999, Hinton 2001: 73, Scheffler 2003: 22, Daniels 2003, and Munoz-Dardé 2005). Where the two positions differ is with respect to the second claim stated by these passages: i.e., that realising that a fundamental egalitarian aim requires our distributions to comply with the luck/choice principle.

In fact, scholars have developed a battery of objections against luck-egalitarians' second claim. Let me just briefly mention three of the most salient ones. On the one hand, Elizabeth Anderson (see Anderson 1999: 303-307) has famously shown how the strict realisation of the luck/choice principle in our distributions would command eschewing urgent help to citizens whose imprudent voluntary choices have resulted in precarious results for them (see also Scheffler 2005a: 15). Call this critique *Anderson's objection*. On the other hand, Jonathan Wolff (see Wolff 1998: 109-112) has shown how the strict realisation of the luck/choice principle in our distributions would require from citizens to unveil shameful aspects about themselves regarding their lack of talent. Call this critique *Wolff's objection*.

Add to these influential objections against luck-egalitarians' second claim what I call the *dysfunctional market objection*. G.A. Cohen himself poses this objection. He states:

[O]ne man's choice is another man's luck [...] Choices both to give and to buy have the property that it is accidental who is favored by them: You and I offer a commodity at £10, and it is an accident from whom a purchaser decides to buy, even in the most "perfect" of markets. And the underlying point might be that a luck egalitarian can't allow for *any* transactions. Sure, she can allow transactions that preserve absence of luck in the distribution, but that won't confer much choice. (Cohen 2011a: 143)

Let us suppose that the purchaser randomly decides to get the commodity from Cohen's store and not from mine. The *dysfunctional market objection* states that if the distribution of inequalities between Cohen and I is to follow the luck/choice principle, then either I should be compensated by the inequality that the purchaser's choice has generated or her purchase from Cohen's store should be impermissible. This is so because it would only have been bad brute-luck for me that the purchaser decided to buy at Cohen's and not at my store. Accordingly, to preserve an egalitarian distribution according to the luck/choice principle, one of the two mentioned paths must be taken (i.e., compensating me for the inequality or making arbitrary market choice impermissible). Now replicate this simple example to all market transactions for, as Anderson 2007 argues and Cohen himself acknowledges in the passage quoted, in every market, no matter how perfect it is, citizens make arbitrary choices regarding where to buy their commodities. In this sense, there will always be an unavoidable element of luck affecting third parties in every single market transaction. In similar fashion, Lazenby 2010

defends the view that citizens always make arbitrary choices in their acts of offering 'material' and 'immaterial' gifts to one another. Variations on persons' circumstances because of such gifts are unavoidable. While the circumstances of some are improved the circumstances of others deteriorate. There is, again, an unavoidable element of luck in every act of giving. If the luck/choice principle were to inform our distributions, then either random giving acts to second parties should be followed by compensation to affected third parties, or they should be impermissible. For, in accordance with Cohen's 2011a: 143 dictum, we can say: one person's choice to buy or to give is at once *always* another person's good brute-luck (the one selling the commodity or receiving the gift) and another person's bad brute-luck (the one not selling the commodity or not receiving the gift).

Now, to these and similar objections (such as the *levelling-down objection* —see Raz 1986: 227-231 and Parfit 1997: 211) luck-egalitarians typically reply by stating that egalitarian justice, as understood by them, is but one aspect that is to be considered when offering principles for the regulation of social affairs in the real world. So it may be that the achievement of perfect egalitarian justice, in accordance with the luck/choice principle, in effect implies the outcomes foreseen by the *levelling-down objection*, *Anderson's objection*, *Wolff's objection* and the *dysfunctional market objection*; however, since equality is but one principle that is to inform material distributions in a given society, it may be limited and complemented by other principles, thereby ensuring that such outcomes are avoided. Cohen,

for instance, explicitly endorses this reply. For he defends what he calls a 'weak-*equalisandum* claim'; that is, a claim stating that people 'should be as equal as possible in some dimension [welfare, resources, capabilities, flourishing] but subject to whatever limitations need to be imposed in deference to other values' (Cohen 2011b: 5) —see Parfit 1997: 211ff.³¹

Cohen has made of this reply a debate on the very nature of first principles of justice. In § 5.1 I will consider his ideas in this regard in the context of his criticism of what he calls 'the constructivist approach to social justice' (Cohen 2008: 274). For now, I only want to point out that this successful reply in the context at stake comes with an incredibly high price for luck-egalitarians. For the reply implies that luck-egalitarians do not aspire to a fully *egalitarian* society in their own terms, since such a society would observe the consequences foreseen by the *levelling down objection*, *Anderson's objection*, *Wolff's objection* and the *no functional market objection*.

Accordingly, *pace* Kymlicka 2002: 86 and Tan 2008: 665, luck-egalitarianism *cannot* be seen as a position describing what is needed of a distribution that realises the fundamental egalitarian aim (i.e., showing respect for citizens' *moral equality*). In fact, as the objections show, realising such an aim requires limiting the influence of the luck/choice principle in determining the outcome of our distributions. If luck-egalitarians were to

³¹ Perhaps the best example of this reply is Larry Temkin's famous passage: 'I, for one, think that inequality is bad. But do I *really* think that there is some respect on which a world where only some are blind is worse than one where all are? Yes. Does this means that I think it would be better if we blinded everybody? No. Equality is not all that matters. (Temkin 1993: 282. Italics in the original)

insist that their position *is* one about what is needed to realise the fundamental egalitarian aim, this would come at the cost of holding that we should aim at respecting citizens' moral equality *imperfectly*. However, this looks like a rather wrong ideal to pursue. I call this the *wrong ideal objection*, and it is prompted by the typical luck-egalitarian response to avoid other objections.³²

Finally, notice how Rawls's position cannot be charged with the *wrong ideal objection*. Kymlicka is mistaken in holding

...what it would be for a distributive scheme to fulfil the basic aims of Rawlsian theory: a distributive scheme that respects the moral equality of persons by compensating for unequal circumstances while holding individuals responsible for their choices. (Kymlicka 2002: 86)

These interpretative remarks fly on the face of Rawls's own discussion on what is needed to protect the moral equality of persons. Rawls's explicit discussion on this matter is prompted by an objection to a distribution of material wealth guided by the difference principle: namely, that the worse-off of society may see their equal status affected by the inequalities that the difference principle allows. To this objection, Rawls responds:

Of course, it does not follow that in a just society everyone is unconcerned with matters of status. The account of self-respect as perhaps the most important primary good has stressed the great significance of how we think others value us. But in a well-ordered society the need for status is met by the public recognition of just institutions, together with the full and diverse internal life of the many free communities of interest that equal liberties allow. The basis for self-respect in a just society *is not then one's income share but the publicly affirmed distribution of fundamental rights and liberties. And this distribution*

³² A similar claim is made by Wolff in a paper revisiting his seminal critique: '[E]ven for those who believe luck egalitarianism to be ideal theory, there is still a great deal of thinking to do about whether we should aim to implement it in the real world, whether there is a path from there; what the steps are in the path [...] The case of shameful revelation can be taken as an illustration of how the process could go wrong, and to point out the need of broader reflection'. (Wolff 2010: 347)

being equal, everyone has a similar and secure status when they meet to conduct the common affairs of the wider society. (TJ: 477. Italics added)

In this passage Rawls holds that citizens in a well-ordered society do not ground their equal status to one another on the putative parity of their material incomes. Rather, they ground it in the fact that they enjoy equal access to rights and liberties, and thus in the fact that they enjoy what Rawls (TJ: 476) calls *equal citizenship*.

The contrast identified in Rawls's passage is reproduced in the debate we have just evaluated between luck-egalitarians and social-egalitarians regarding what is required to realise the fundamental aim of egalitarian justice. Against luck-egalitarians' second claim (i.e., that realising the fundamental egalitarian aim requires our distributions to comply with the luck/choice principle), social-egalitarians defend variations of the idea that our distributions should be free of *oppression*-determinants, so that regardless of the material inequalities arising by luck and choices within such distributions, we nevertheless relate to each other as equal persons on equal footing within social and political structures (see Anderson 1999, 2007, 2008, Scheffler 2003: 22, Hinton 2001: 73, Munoz-Dardé 2005). According to Rawls, the equal distribution of liberties and rights ensures that.³³

I have defended the claim that it is a fundamental task of political philosophy to offer descriptions of *how things should be*; that is, *ideals* to arrive at in our actual social proceedings. The importance of this task

³³ Daniels 2003: 247-248 and Scheffler 2003: 8-13 offer similar interpretations of Rawls by highlighting other aspects of his theory.

discredits *the non-use critique*. Also, we saw that from the point of view of egalitarian justice, the *ideal* of a society in which citizens' moral equality is fully respected is particularly important. Rawls's realistic utopia of a well-ordered society seems to be a better description of such an *ideal* than a society in which distributive justice is fully informed by the luck/choice principle. However, as we have seen in the previous chapter, such a description might need to undergo some modifications to allow for full rectification of historical wrongs.

4.2 The *Institutional Injustice Critique* and the Two Principles

There is another way in which one can complain about Rawls's theoretical silence regarding gender and racial relations: to wit, the idea that Rawls's lack of concern for race and gender discrimination obscures the fact that in a well-ordered society—that is, a society perfectly regulated by justice-as-fairness— injustices regarding racial or gender discrimination will arise within society's institutional structures. Call this the *institutional injustice critique*.

First of all, notice the similarity between the *institutional injustice critique* and mine. Both hold that the existence of certain injustices is compatible with Rawls's original description of a well-ordered society. Notice also that the *institutional injustice critique* is usually dismissed by the same consideration with which I started my inquiry in § 2.1: i.e., the ideal

character (in the technical sense) of Rawls's theory. Consider, for instance, what Seana Shiffrin reports regarding Rawls's treatment of racial relations:

The problem of race, for Rawls, may be a problem that arises in non-ideal theory, not ideal theory [...] This is an appealing thought and has some plausibility. Certainly, many of the pressing issues regarding race, such as reparations and affirmative action, are intimately connected to redress for and reconstruction in the face of public failures and wrongs toward people of color [...] If the two principles were implemented, the significance of class distinctions as we know them would alter dramatically. Even if class distinctions remained, the absolute level of social stratification would reduce [...] While these motivations may explain the omission of an in-depth discussion of race by Rawls, I do not think they fully justify the omission of explicit anti-discrimination principles. (Shiffrin 2004: 1654)

Now consider what Lisa Schwartzman comments regarding Rawls's treatment of gender relations:

On a similar note, a Rawlsian might argue that because a *Theory of Justice* deals mainly with ideal theory, that is, with a world in which there is 'full compliance' (as opposed to 'partial compliance') with a theory of justice, Rawls need not address the specific forms of injustice based on factors such as gender and race. There are several problems with this line of argument, however. (Schwartzman 2006: 68)

At this stage of contemporary political philosophy, it is curious how often the ideal character of Rawls's theory is pointed out as an instantaneous way of dismissing the *institutional injustice critique*; especially considering that the social understanding of egalitarian justice is now more prominent (with its strong emphasis on the importance of avoiding *oppression*-determinants in our distributions), and also how frequently it is defended that Rawls's theory offers a defence of such an understanding (see Anderson 1999, Daniels 2003, Scheffler 2003, Freeman 2007b). I therefore share the dissatisfaction of the quoted scholars. Also curious is the fact that Kymlicka's critique highlighting Rawls's absence of concern for the relations between majority and minority

cultures has not been dismissed in such a way (as I reported at the end of § 2.1).

All the same, I think that the *institutional injustice critique* only has real force when framed as criticism of the lack of concern for historical injustice of Rawls's theory, and not as a way of defending that *new* injustices would arise in a well-ordered society. For, aside from the possible absence of historical rectification in cases such as **scenarios 1, 2 and 3** (see § 3.1), it seems hard to envision that the correct regulation of the principles of justice-as-fairness would allow the irruption of new forms of injustice or discrimination in given society. Contrarily, Shiffrin holds:

To put it concretely, it is unclear what specific provision of the two principles would directly condemn as unjust the treatment of Rosa Parks and countless other African-Americans who were told they had to sit at the back of the bus. (Shiffrin 2004: 1647)

Note that Shiffrin is not referring to the injustices related to the Jim Crow laws as past injustices in need of rectification. Rather, her claim is that the *possibility* of this type of discriminatory policies is not directly condemned as unjust by the principles of justice-as-fairness. Thus, her concern is about how the forward-looking principles of justice-as-fairness ensure that such policies would not be enacted in a well-ordered society.

Shiffrin considers which basic liberty according to Rawls would be violated by the mentioned discriminatory policies. She concludes that in Rawls's framework these policies command something close to a violation of the 'freedom from arbitrary arrest and seizure as defined by the rule of law' (*TJ*: 53). While Shiffrin concedes that 'arbitrary treatment arguably conflicts

with the principles underlying the rule of law' [...] 'it is certainly regrettable that the antidiscrimination character of these principles [Rawls's principles] is so submerged' (Shiffrin 2004: 1647).

But it could be argued that the anti-discrimination character of Rawls's first principle (for short, EL) is not submerged. Aside from the liberties of the rule of law on which Shiffrin focuses, the discrimination character of the Jim Crow laws is rejected by EL because of the fact that liberties must be *equally* distributed amongst citizens. In the particular case at stake, while it could be defended (perhaps with some controversy) that freedom of movement is not equally distributed, it is nevertheless uncontroversial that, by requiring some citizens to sit at the back of the bus in function of their race, what Rawls (*TJ*: 53) calls the 'freedom of the person, which includes freedom from psychological oppression' (see also *PL*, 291, *JF*: 44), is been distributed unequally.

It is certainly the case that the importance of this crucial liberty is not fully explored by Rawls. However, consider Freeman's characterisation of such a liberty in the context of introducing all basic liberties:

He [Rawls] mentions five sets of basic liberties: freedom of conscience and freedom of thought; freedom of association; equal political liberties; the rights and liberties that protect the integrity and freedom of the person (including freedom of occupation and choice of careers and a right to personal property); and finally the rights and liberties of the rule of law. *Protections for the physical and psychological integrity and freedom of the person are the most obvious basic rights and liberties, for they forbid unjustified violence, coercion, and enslavement of persons (amongst other things)*. Any reasonable conception of justice, liberal or non-liberal, recognizes these as morally protected rights. (Freeman 2007a: 46. References omitted)

As Freeman suggests in this passage, the rights and liberties protecting persons' physical and psychological integrity are extremely important from the point of view of liberal justice. In this sense, the connection of the freedom from psychological oppression with the social bases of self-respect and the fulfilment of the fundamental egalitarian aim (respect for citizens' moral equality) should be explored more deeply. I think, nevertheless, that appealing to such a liberty within the Rawlsian framework suffices to condemn straightforwardly as unjust the arbitrary treatment enacted by a discriminatory system of law.

Appealing to this basic liberty also clarifies how to answer another of Shiffrin's claims. She argues that bus transportation and other institutional settings in which discrimination might occur, as the paradigmatic case of Rosa Parks illustrates, are not even regulated by the liberties ensuring the rule of law:

However broadly understood, the guarantee of the rule of law only applies to the state and its legal system. Importantly, it does not represent the idea that other forms of discrimination in the provision of public amenities by non-governmental actors are unjust in addition to immoral. Many bus companies, lunch counters, and hotels, are, after all, privately operated. (Shiffrin 2004: 1648. References omitted)

It may be a matter of debate whether the institutions of public transportation are directly regulated by the principles of justice as part of the basic structure of society. It is nevertheless clear that such institutions are at least indirectly regulated by them, just as other private associations are. To see this, recall the divide between *local* and *domestic* justice adopted by Rawls (*JF*: 10ff). On the one hand, associations such as churches and universities are directly

regulated by diverse principles of local justice. These, for instance, may not command strict democratic policies in the structural organisation of such associations. On the other hand, domestic principles (such as those of justice-as-fairness) have a twofold role: first, they are to regulate directly the institutions of the basic structure and, second, they are to regulate indirectly local associations and their principles:

In general, principles for the basic structure constrain (or limit), but not determine uniquely, the suitable principles of local justice. (*JF*: 11-12)

As citizens we have reasons to impose the constraints specified by the political principles of justice on associations; while as members of associations we have reasons for limiting those constraints so that they leave room for a free and flourishing internal life appropriate to the association in question. (*JF*: 165)

Rawls holds that the principles of justice-as-fairness regulate indirectly those of local justice by putting constraints on them. This restriction suffices to ensure that even if the institutions providing public transportation are not part of the basic structure (and thus reassemble more the status of hotels, clubs, universities, and churches), they nevertheless *cannot* adopt a policy principle commanding preferential treatment to some citizens considering factors such as race, since the liberties protecting persons' integrity would be impaired.

It seems, then, that EL as it stands can be straightforwardly understood as a principle against discrimination. Now, even if it is shown that EL is a straightforward principle against discrimination, it is not so clear that it directly commands historical rectification —that is, it is not so clear

that it can be interpreted as a principle of historical rectification. Consider, again, what Shiffrin says in the passage with which I opened this section:

The problem of race, for Rawls, may be a problem that arises in non-ideal theory, not ideal theory [...] This is an appealing thought and has some plausibility. Certainly, many of the pressing issues regarding race, such as reparations and affirmative action, are intimately connected to redress for and reconstruction in the face of public failures and wrongs toward people of color [...] While these motivations may explain the omission of an in-depth discussion of race by Rawls, I do not think they fully justify the omission of explicit anti-discrimination principles. (Shiffrin 2004: 1654)

I have tried to show that Shiffrin is mistaken in arguing that EL is not an explicit anti-discrimination principle. However, I agree with her on the following idea:

It does seem strange to classify, implicitly, anti-discrimination principles along with principles of redress. Anti-discrimination principles seem more analogous to the other, already-acknowledge basic liberties: they demarcate a standard of treatment that is *forward-looking* and that aims at regulate a latent or explicit hazard. (Shiffrin 2004: 1655. Italics added)

Shiffrin is right in arguing that EL has an obvious forward-looking character. In fact, this is one of the main reasons why I think that justice-as-fairness could be complemented by a backward-looking principle. For, while it is certainly true that we could appeal to the very freedom from psychological oppression as a way of defending the need of historical rectification in a well-ordered society, appealing exclusively to such a liberty for that purpose might place too much explanatory weight on it. For instance, the need to adopt a critical historical narrative (one of the chief rectificatory measures in my view) might not follow directly from the need to ensure citizens' freedom from psychological oppression. Note, nevertheless, that I am not suggesting that there is no plausible way of arguing in favour of full historical

rectification by appealing to EL. However, as I said in § 3.4, it seems at least theoretically convenient to include a specific backward-looking principle in order to express direct commitment to that goal. This principle, unlike EL, would be a straightforward backward-looking principle —or, as Shiffrin holds, a ‘principle of redress’.³⁴ As I said in § 3.2, such a principle could bring theoretical balance to justice-as-fairness, for the importance of fulfilling the prerogatives of justice in relation to the past, present, and future of a well-ordered society would be explicitly addressed. Also, adding a specific backward-looking principle could help to make explicit the commitment of Rawls’s theory to the idea that a liberal society is not fully just until historical rectification has taken place —so that ‘closing the books’ or drawing a ‘thick line’ through the past is never an option from the point of view of justice. These reasons ground the pertinence of including a backward-looking principle as part of an *ideal* liberal theory of justice.³⁵

In sum, while EL explicitly establishes the commitment of justice-as-fairness to anti-discrimination, it does not explicitly establish a commitment to the need for historical rectification. Therefore, while there is reason to add a specific backward-looking principle to such a conception of justice, there seems to be no reason to include a further anti-discrimination principle within it.

³⁴ This use differs from what Rawls (*TJ*: 86) calls the ‘principle of redress’.

³⁵ Remember Simmons’s revealing remark quoted in § 3.2: ‘A quite different content to ideal theory (for instance, one that was more sensitive to the need to redress historical injustice) could be defended while still subscribing wholeheartedly to the Rawlsian version of the ideal-nonideal distinction’ (Simmons 2010: 32-33).

Let me now turn to the relation between discrimination and the fair equality of opportunity principle (for short, FEO). Of course, many discriminatory acts would not occur in a well-ordered society *because* of the regulation of FEO. Shiffrin 2004: 1650 is very aware of this fact: 'employment discrimination is obviously incompatible with the principle of fair equality of opportunity'. In fact, partly because of this she offers powerful arguments against the priority of EL over FEO (see Shiffrin 2004: 1672). She notes that Rawls's main argument in favour of the priority of EL over the second principle only applies to the difference principle, for Rawls holds that the second principle as a whole should not be part of the constitutional essentials because 'the history of successful constitutions suggests that principles to regulate economic and social inequalities [...] are generally not suitable as constitutional restrictions' (PL: 337). As is clear, this defence of the priority of EL assumes that FEO's main function is helping in the correct regulation of the economy—in the very same way as that of the difference principle.

However, this interpretation of FEO clashes with Rawls's own defence of the priority of FEO over the difference principle:

I have not maintained that offices must be open if in fact everyone is to benefit from an arrangement. For it may be possible to improve everyone's situation by assigning certain powers and benefits to positions despite the fact that certain groups are excluded from them [...] But the principle of open positions [FEO] forbids this. It expresses the conviction that if some places were not open on a fair basis to all, those kept out would be right in feeling unjustly treated even though they benefit from the greater efforts who were allowed to hold them. (TJ: 84)

This passage—which establishes the priority of FEO over the difference principle—explicitly states that FEO is *not* a principle concerned with the

efficient performance of the economy. For even if unfair and unequal opportunities to members of different groups were to benefit the worse-off, such a differential treatment would simply be unjust. Therefore, the basis for the priority of FEO over the difference principle shows that the basis for the priority of EL only applies to the difference principle.

Because of this reason, it could be argued (along Shiffrin's lines) that, in a well-ordered society, citizens ground their equal status not only in the fact that they possess equal rights and liberties —as the Rawlsian passage reviewed at the end of § 4.1 establishes (see *TJ*: 478)—, but also in the fact that they possess fair opportunity. In other words, Rawlsian *equal citizenship* encompasses fair opportunity. Rawls's own defence of the priority of FEO over the difference principle seems to support this idea.

Having said that, we should consider the main reason why Shiffrin rejects the claim that FEO can be understood primarily as an antidiscrimination principle:

Deriving principles of anti-discrimination from the fair equality of opportunity principle is an unsatisfying route of ensuring racial equality and of marking the wrong of racial discrimination [...] Although a crucial reason that discrimination is so offensive to our sense of justice has to do with the distorting effect on individuals' career prospects, other opportunities and equal access to power, it is not the exhaustive, or perhaps the central, objection to racial discrimination. (Shiffrin 2004: 1650)

Shiffrin's point here is that, even when many discriminatory acts would not occur because of the regulation of FEO, the commitment to avoid discrimination is more fundamental than the need to secure fair and equal opportunities across the citizenry. So we are back to the idea that the only

direct anti-discrimination principle within justice-as-fairness is found in EL: what it is wrong with discrimination is that it shatters the personal integrity of the discriminated person. The fact that her competitive chances can be affected seems to be a by-product of that wrong.

In fact, I made a parallel remark in § 3.1 regarding the mitigating effect of FEO on the effects of historical injustice: i.e., that the normative significance of rectifying past political violence is more fundamental than the need to provide fair competition for social positions of advantage across the citizenry—even if in a well-ordered society most of the effects of historical wrongs would have been mitigated *because* of the regulation of FEO. Rather, the normative significance of setting forth rectificatory measures is fundamentally connected to the need of acknowledging institutional past wrongdoing as such, as something that should not have occurred, and as something that now society officially regrets. Furthermore (as we also saw in § 3.1), just as with EL, some mechanisms of historical rectification do not seem to follow directly from the correct regulation of FEO.

In any event, note that Shiffrin's version of the *institutional injustice critique* does not warrant the inclusion of an explicit anti-discrimination principle to Rawls's original theory.

4.3 The *Non-Institutional Injustice Critique* and the Egalitarian Ethos

So far we have seen that both the *no-use critique* and the *institutional injustice critique* fail against Rawls's theory. But there is a third way in which one can

complain about Rawls's theoretical silence regarding gender and racial relations: to wit, the idea that Rawls's lack of concern for race and gender discrimination obscures the fact that in a well-ordered society —that is, a society perfectly regulated by justice-as-fairness— injustices regarding racial or gender discrimination would arise in *non-institutional* settings. Call this the *non-institutional injustice critique*.

This critique has some initial plausibility. As we have seen, Rawls's principles guarantee non-discrimination in local and domestic institutions. However, citizens also interact outside those domains: they go to the park, walk on the street, settle in certain neighbourhoods, talk to each other, make decisions regarding who they want to befriend, etc. Considering this, even when domestic and local institutions would not allow acts of discrimination when the principles of justice-as-fairness are in place, subtle forms of discrimination might still occur in this further domain. Think, for instance, of the possibility that even if public transportation is strictly regulated by a general policy prohibiting discrimination, some particular citizens may still deny them seats, or change places in order to avoid sharing a seat, or try to persuade others to sit somewhere else by looking at them disrespectfully or aggressively because of their race, gender, or sexual orientation. Kymlicka neatly illustrates the point:

[L]egal requirements of non-discrimination have increasingly been applied to 'private' firms and associations. This extension of non-discrimination from government to civil society is not just a shift in the scale of liberal norms, it also involves a radical extension of the obligations of liberal citizenship. For the obligation to treat people as equal citizens now applies to the most common everyday decisions of individuals [...] Liberal citizens must learn to interact in

every day life settings on an equal basis with people for whom they might harbour prejudice. (Kymlicka 2002: 301)

The discussion of the correct regulation for this domain of citizens' interaction then takes us to a debate regarding the adequate *egalitarian ethos* that should govern citizens' personal lives.

Most of the discussion on such an ethos has focused on whether the difference principle should inform citizens' personal job-market decisions and how that might contradict Rawls's argument in favour of economic incentives (see Cohen 1992, Williams 1998, Van Parijs 2003, Scheffler 2006, Titelbaum 2008, Shiffrin 2010). We can distinguish three positions in this debate. On the one hand, Cohen extensively argues that the citizens of a well-ordered society should guide their personal job-market decisions in accordance with the difference principle, so that they would not ask for economic incentives to work harder, since this would make the worse-off as well-off as possible (see Cohen 1992, 1997). On the other hand, some scholars deny that citizens should directly guide their personal job-market decisions by the difference principle (see Williams 2008 and Scheffler 2005, 2006). Scheffler, in particular, supports this claim by appealing to a division of moral labour between principles for the basic structure required in securing background justice and principles for personal conduct (see Scheffler 2005, 2006). Finally, scholars in the third position have argued that, while a 'correlate' of the difference principle must indeed regulate the personal conduct of citizens in a well-ordered society, so does a 'correlate' of the first principle of justice. Titelbaum, in particular, argues that a 'full ethos'

including correlates of both principles would allow citizens to exercise what he calls 'productive latitude' (i.e., not always trying to make the worse-off as well-off as possible when taking into account job-market decisions), as granted by the priority ascribed to the correlate of the first principle over the correlate of the second one within the full ethos (see Titelbaum 2008: 315-322).³⁶

All this work on the putative egalitarian ethos appropriate for a well-ordered society has been dominated by Cohen's original critique of Rawls's incentives argument. However, considering the fundamental egalitarian aim—i.e., respect for citizens' moral equality—the relevance of such an ethos is much more patent in light of the *non-institutional injustice critique*. By focusing on how the two original principles of justice would properly inform the stated ethos, these authors have failed to see that, more generally, citizens should also guide their interactions by a norm ensuring that they would respect the moral equality of all. Ascribing such content to the ethos is less controversial and, in this sense, independent of which of the three mentioned positions prompted by Cohen's critique of Rawls's incentives argument is correct; it is, also, a more urgent task in light of the fundamental egalitarian aim.

³⁶ Cohen also defends the possibility of exercising 'productive latitude'. He does so by appealing to what he calls a legitimate 'agent-centered prerogative' (see Cohen 1992: 302). However, as Titelbaum 2008: 322 points out, Cohen does not offer an argument as how such a prerogative would be justified to the worse-off. This, Titelbaum holds, is a further advantage of defending a Rawlsian full ethos of justice as his.

There are four main Rawlsian theses grounding the existence of such a norm within the ethos governing citizens' personal interactions in a well-ordered society. First, and most importantly, the fact that the parties in the original position agree to a principle of *mutual respect*:

This is the duty to show a person the respect which is due to him as a moral being, that is, as a being with a sense of justice and a conception of the good [...] Mutual respect is shown in several ways: in our willingness to see the situation of others from their point of view, from the perspective of their conception of their good, and in our being prepared to give reasons for our actions whenever the interest of others are materially affected. (*TJ*: 297)

According to Rawls, a duty of mutual respect so defined would be agreed to in the original position. This means that the parties consider that in a well-ordered society all citizens would observe this duty in their personal interactions. In direct response to the *non-institutional injustice critique*, we should notice that 'not to have or act on racist attitudes [...] follows from the duty of mutual respect' as Alex Voorhoeve 2005: 5 has pointed out.

A second Rawlsian thesis in this regard is that citizens accept all the main ideas of the shared political culture of a liberal society. One of such ideas, the most important one —as Rawls (*JF*: 5) states—, is the idea that a liberal society is a fair system of social cooperation between free and *equal* citizens. Citizens are all supposed to accept this fundamental idea within a well-ordered society, and one thing they have to do to honour it is, precisely, to respect their co-citizens' moral equality in both formal and informal social settings.

A third Rawlsian thesis in this regard is that in a well-ordered society what Rawls calls the *political virtues* would shape citizens' moral character.

These virtues include, 'the virtues of social cooperation such as civility and tolerance, of reasonableness and the sense of fairness' (PL: 194). Were these virtues to shape citizens' moral character, it is hard to envision that they would commit discriminatory acts.

Finally, a fourth Rawlsian thesis is that citizens would observe what Rawls (PL: 217) calls the *duty of civility*: to wit, only offer reasons in political matters that one sincerely thinks are acceptable from the point of view of all reasonable comprehensive doctrines. Now, while the limits of this duty are not clear (see PL: 253-254), it is nevertheless clear that it applies to citizens' individual political actions and not only to those of officials, legislators and judges (see PL: 217).

These four theses show that the egalitarian ethos of a well-ordered liberal society would be informed by a norm commanding the respect for citizens' moral equality. However, Rawls's original rationale in favour of the first thesis must be revised in the context of the ideas of political liberalism. For Rawls seems to have offered the wrong reasons in support of the claim that the parties in the original position would agree to the duty of mutual respect (see TJ: 297). Admittedly, Rawls's discussion of this matter is brief:

Now our self-respect normally depends upon the respect of others. Unless we feel that our endeavors are respected by them, it is difficult if not impossible for us to maintain the conviction that our ends are worth advancing. Hence for this reason the parties would accept the natural duty of mutual respect. (TJ: 156)

This passage is a statement of what Bird 2008: 17 (as we saw in § 1.2) calls 'the dependency thesis' —i.e., the claim that all that deficits in mutual respect bring about deficits in self-respect. I have argued against such a thesis. Yet

other passages evince that Rawls thinks that the parties would agree to the duty of mutual respect out of self-interested reasons:

Now the reason why this duty would be acknowledged is that although the parties in the original position take no interest in each other's interests, they know that in society they need to be assured by the esteem of others their associates. Their self-respect and their confidence in the value of their own system of ends cannot withstand the indifference much less the contempt of the others. *Everyone benefits from living in a society where the duty of mutual respect is honoured. The cost of self-interest is minor in comparison with the support for the sense of one's own worth.* (TJ: 297. Italics added)

Here Rawls holds that the parties would elect the duty of mutual respect out of preoccupation for the possible psychological and moral consequences of not being respected in turn. One might initially think that this reasoning fashions an unproblematic impulse for reciprocity in social interactions. However, let us suppose that citizens could indeed develop successfully their own system of ends in spite of the indifference and disrespect of others. Does this mean that in such situations, considering their incredible psychological strength and self-sufficiency, citizens are no longer obliged to show respect for their co-citizens? The answer seems to be, plainly, no. However, if the reason why citizens respect others is merely because of the importance that being respected in turn holds for the pursing of their way of life, the outcome just outlined seems to follow. The problem is, then, that Rawls's original reason in support of the duty of mutual respect makes its observance by citizens in a well-ordered society a *modus vivendi*.³⁷

³⁷ This original rationale in favour of the duty of mutual respect is more akin to a rational choice reading of the original position in *A Theory of Justice*. However, as we have seen in § 2.2, Rawls was quite clear in his later work regarding his dissatisfaction with this reading of the original position — see (JF: 82n).

Setting aside this modification to Rawls's main rationale behind the duty of mutual respect, notice three things. First, that Rawls's theory has all the ideas needed to justify that an egalitarian ethos would regulate citizens' personal conduct in a well-ordered society (as opposed to what Titelbaum 2008 holds in the context of Cohen's critique). Second, that such an ethos does not violate justice-as-fairness's non-comprehensive character (as opposed to what Van Parijs 2003 holds in the context of Cohen's critique). Third, that such an ethos does not violate the moral division of labour in Rawls's theory (as opposed to what Scheffler 2005 holds in the context of Cohen's critique).

If a norm commanding respect for citizens' moral equality properly informs the egalitarian ethos of a well-ordered society, both the *institutional injustice critique* and the *non-institutional injustice critique* are fully met. For it would have been shown that no discriminatory acts would take place in the domestic, local, and personal domains. Note also how the inclusion of an egalitarian ethos in the description of a well-ordered society helps to respond to more feminist-oriented versions of such critiques. The final statement of Rawls's theory seems to favour the claim that the institution of the family is part of the basic structure and thus regulated by Rawls's original principles of justice (see *JF*: 162-66). If we add to this the fact that citizens would observe the egalitarian ethos, *pace* Arneson's 1999 critique of FEO, the social and cultural roles assigning different labour burdens and carrier expectations to women and men within the family would disappear. So long

as such roles are incompatible with the respect of women's moral equality, citizens would not support them once they embraced the egalitarian ethos.

Notice, then, that a Rawlsian response to the *institutional injustice critique* and the *non-institutional injustice critique* does not require the inclusion of further principles to justice-as-fairness. Rather, it requires working out the proper ethos that would govern the private relations of citizens within a well-ordered society. This conclusion marks a sharp contrast with the claims I made in § 3.1—§ 3.2. I argued that without the addition of a further principle for the regulation of the basic structure, full historical rectification might not be guaranteed in a well-ordered society. For all *retribution*, *compensation* and *recognition*-driven measures needed to rectify historical wrongs occur within and with support of basic institutional settings. Conversely, it is difficult to see what particular principle regarding historical rectification must inform the egalitarian ethos. So long as such an ethos is guided by a general norm as stated above, the interaction of victims of past wrongs with the rest of the citizenry in private and informal settings is guaranteed to be a respectful one.

4.4 Concluding Remarks

This concludes my assessment of some of the critiques against Rawls's theoretical silence regarding racial and gender relations. As we have seen, aside from the need to flesh out the proper egalitarian ethos of a well-ordered society (by pointing at the fact that it includes a norm commanding a

commitment to the fundamental egalitarian aim —i.e., respect for citizens' moral equality), no further addition to justice-as-fairness is needed to meet the revised critiques. In the following chapter I will resume my assessment of the discussion about the nature of egalitarian justice by exploring Cohen's critique of Rawls's concept of justice.

5

Practice-Dependence, Constructivism, and Liberalism

In the previous chapter we saw that in contemporary political philosophy there is a wide dispute about the correct understanding of the nature of egalitarian justice. As with many disagreements in current political philosophy, this one starts as a criticism to Rawls's ideas in *A Theory of Justice*. Perhaps the most developed contemporary critique in this regard is presented G. A. Cohen in his book informatively called *Rescuing Justice and Equality* (see Cohen 2008). In this chapter, I will review his critique of what he calls the 'constructivist understanding of justice' and also how the constructivist methodology relates to political liberalism. In § 5.1 I defend three things: first, that Cohen is mistaken in presenting his critique of Rawls's concept of justice as a critique of Rawls's constructivism. Second, that Cohen's first condition of fundamental principles of justice (stating that such principles must be fact-insensitive) is trivial. Third, that Cohen's second condition of fundamental principles of justice (stating that such principles must be derived from considerations of justice alone) creates what I call *Cohen's dilemma*. In § 5.2 I argue that only a meta-ethical understanding of

constructivism, as opposed to one offered at a normative justificatory level, must be considered a constructivist view proper. Finally, in § 5.3 I hold that justice-as-fairness, in order to serve the purposes of political liberalism, is better interpreted as uncommitted to a constructivist view so understood. This forces the elimination of Rawls's mirroring ideas of *full autonomy* and *political constructivism* and, ultimately, prompts a deflationist interpretation of the original position.

5.1 A Concept of *Pure* Justice VS a Concept of *Impure* Justice

Throughout this thesis I have endorsed a particular understanding of how to do political philosophy that is now sometimes called *practice-dependent* or *practice-based* (see James 2005: 282, Mackled-Garcia 2008: 246, Sangiovanni 2008: 2, Valentini 2011: 399-400). What distinguishes this understanding is the claim that facts about a given practice (e.g., local, domestic or global interactions) must be seriously considered when justifying principles of justice for such a practice.

This way of doing political philosophy is, of course, widely inspired by Rawls's theory. A commitment to such a crucial claim is stated in the very first pages of *A Theory of Justice*: 'the correct regulative principle for anything depends on the nature of that thing' (*TJ*: 25). Later on, Rawls affirms: 'conceptions of justice must be justified by the conditions of life as we know it or not at all' (*TJ*: 398). As we have seen in § 2.1, Rawls holds that, when proposing domestic principles of justice, we must consider all available

knowledge about human psychological regularities as well as the general social facts about liberal societies (see *TJ*: 119). In fact, I have partly pressed Rawls's theory regarding the importance of historical rectification precisely on these terms (see §1.1-§ 2.1).

The practice-dependent approach just outlined, as represented by Rawls's work itself, has been recently criticised by G. A. Cohen (see Cohen 2008). In this section I advance three objections to such a criticism. First, I defend that while Cohen claims that he is criticising Rawls's constructivism, his critique exclusively targets Rawls's understanding of the concept of justice. In order to show this, I hold, there is no need to engage with Cohen's critique itself. Second, I argue that the first part of Cohen's critique (stated by the condition that fundamental principles of justice must be fact-insensitive) is trivial. Third, I hold that the second part of Cohen's critique (stated by the condition that fundamental principles of justice must not be derived from non-justice considerations) speaks against Cohen's own preferred principle of justice, creating what I call *Cohen's dilemma*. Let me start with the first of these issues.

Cohen criticises what he calls 'the constructivist approach to social justice', which he defines as follows:

In its most general description, constructivism is the view that a principle gains its normative credentials by being the product of a sound selection procedure. But I am not concerned in this but in entirely general form. I am concerned with, precisely, the constructivist approach to social justice in particular, which is constructivism understood as characterized above, but with respect to fundamental principles of justice in particular, *and* that proceeds by putting and answering the question "What rules of governance are to be adopted for our common social life?" Unless otherwise indicated, that is what I shall mean by 'constructivism' here. (Cohen 2008: 274-275. *Italics in the original*)

Something needs to be immediately noticed about what Cohen takes to be his target. As the passage states, Cohen criticises constructivism as the view holding that:

- (1) The normative credentials of fundamental principles of justice are gained through their being the result of a specified procedure.
- (2) Fundamental principles of justice are the ‘rules of governance [...] to be adopted for our social common life’.

A view committed to (2), Cohen explains at length, identifies fundamental principles of justice (i.e. principles that are not derived from other principles) with the optimal principles of regulation for society (see Cohen 2008: 276). Cohen then claims that a theory committed to (2) is mistaken because identifying fundamental principles of justice with principles of regulation for the institutions of society makes the former sensitive both to facts about human psychology and social regularities, as well as to other important values other than justice alone—for only principles of regulation sensitive to these two things can successfully regulate a society like ours.

Of course, Rawls explicitly champions a position committed to (2): ‘It is a fundamental error of a *Theory of Justice* that it identifies the first principles of justice with the principles that we should adopt to regulate society’ (Cohen 2008: 265). This putative error goes back to Rawls’s (*TJ*: 5) basic distinction between a concept and a conception of justice. According to Rawls, the concept of justice is defined by a problem: what the fair terms of social cooperation are in a society in which every one has similar capacities and cooperates to the same extent. A conception of justice, on the other hand,

is formed by principles to solve this problem. The character of the problem as defined by the concept of justice (finding principles to regulate the institutions of the basic structure of society) commands the careful consideration both of facts about human nature and society, as well as of other important values that must guide the distribution of burdens and benefits of social cooperation such as Pareto optimality, publicity and stability.

As I said, Cohen thinks that while it is indeed correct to consider facts about human nature and society —as well as values other than justice itself— to justify principles of regulation for the institutions of society, it is incorrect to take into consideration such things to justify fundamental principles of justice. Whatever fundamental principles of justice are, Cohen 2008: 281 claims, they must be the result of considerations of justice alone. According to Cohen, then, there are two conditions that fundamental principles of justice must fulfil. Let me call them, respectively, the *fact-insensitivity condition* and the *single-value condition*:

The fact-insensitivity condition: fundamental principles of justice are not justified on the basis that certain empirical facts are the case.³⁸

The single-value condition: fundamental principles of justice are justified by appealing exclusively to considerations dependent on the value of justice.³⁹

³⁸ This is only one way of framing this condition. Several passages in Cohen's text suggest this formulation. Consider, for instance: 'I argue that a principle can respond to (that is, be grounded in) a fact only because it is also a response to a more ultimate principle that is not responding to a fact: accordingly, if principles respond to facts, then the principles at the summit of our conviction, are grounded in no fact at whatsoever' (Cohen 2008: 229).

³⁹ Again, this is only one way of framing this condition. The following passage comes quite close to it: 'I believe that, whatever their content might be, fundamental principles of justice are in no way dependent on the character of any facts, or, indeed, and equally important for

Let me refer to the principles meeting these two conditions as *pure principles* of justice (since they are pure of facts and non-justice considerations) and to the principles that do not meet these two conditions as *impure principles* of justice. Using this terminology, we can say that Cohen's discussion is about whether *impure principles* can be fundamental principles of justice or not. He thinks they cannot. Furthermore, the original position, being a procedure explicitly designed to deliver *impure principles* of justice to regulate the social cooperation of a liberal society, cannot deliver fundamental principles of justice. It may well be that the *impure principles* of justice delivered by the original position are the correct principles for regulating the basic structure of society. However, that is irrelevant to the question of which are the fundamental principles of justice (see Cohen 2008: 291).

Be that as it may, it is striking that Cohen's critique only applies to a view holding (2) —and thus that the principles of regulation of society are fundamental principles of justice. Nothing is said about (1) —the first commitment of a constructivist view that Cohen identifies in his definition of constructivism. In fact, Cohen reiterates that his critique of what he calls constructivism will focus exclusively on the commitment to (2) of such views:

I argue in what follows that the constructivist approach to social justice mischaracterizes justice both because it treats justice as sensitive to certain sorts of facts and because it fails to distinguish between justice and other virtues. The two errors reflect the single disfigurement by constructivism from which I seek to rescue justice, and that is *constructivism's identification of principles of justice with the optimal set of principles to live by, all things considered* [...] *The present chapter is an extended defence of the claim that the*

my purposes, on any consideration of value or principle that are not considerations of justice (Cohen 2008: 281).

constructivist approach to social justice is, for that particular and transparently simple, reason, misguided. (Cohen 2008: 275)

In this passage Cohen himself states that his real and only target is what follows from accepting (2), for notice that the problems that Cohen identifies in a view accepting (2) do not hold for a view exclusively accepting (1). This is because there can be a constructivist view as defined in (1) that simply does not identify principles of regulation with fundamental principles of justice —and, therefore, does not claim that *impure principles* can be fundamental principles.

For instance, let us grant that the luck/choice principle, understood as ‘inequalities are just if and only if certain facts about responsibility obtain with respect to those inequalities’ (Cohen 2008: 300) —or perhaps as inequalities are just if and only if they are due to the voluntary choices of the people affected by them—, is a *pure principle* of justice (i.e., it is fact-insensitive and derived exclusively from considerations of justice). A constructivist committed to (1) would claim that this principle holds its normative credentials because it withstands the appropriate normative procedure (whatever such a procedure happens to be, e.g., an original position, a reasonability test, selection by fully rational creatures, approval from an ideal observer). Contrarily, someone not committed to (1) —a non-constructivist such as Cohen— could claim that this principle holds its normative credentials exclusively from the analysis of our considered

intuitions.⁴⁰ But notice that here the disagreement between the constructivist and the non-constructivist is not about whether this fundamental principle is independent of facts and non-justice considerations. Both agree on that. Rather, the disagreement is about what renders the normative credentials to this fundamental principle so defined.

Likewise, just as there can be a constructivist view (again, if we understand constructivism as in (1)) about *pure principles*, there can also be a non-constructivist view of *impure principles* of justice (i.e., principles that are fact-sensitive and derived from non-justice considerations). The practice-dependent conception of justice set forth by Sangiovanni 2008 is at least compatible with this view. By being practice-dependent, this conception of justice assumes that fundamental principles are *impure*, for all the relevant facts of the practice they are to regulate must be considered when proposing them. However, Sangiovanni remains silent on whether the principles of his conception of justice gain their normative credentials by their being the result of a specified procedure or by the analysis of our considered intuitions (or by any other metaethical or justificatory strategy).⁴¹

Accordingly, because both a constructivist view about *pure principles* of justice and a non-constructivist view of *impure principles* are perfectly

⁴⁰ See Cohen 2008: 4: 'In this conception, the one favoured by many Oxford types like me, we determinate the principles we are willing to endorse through an investigation of our individual normative judgements on particular cases [...] In my philosophically conservative view, that is the only way to go. And when we go that way, we refine, and thereby reach, our deepest normative convictions'.

⁴¹ Notice, however, that other practice-dependent conceptions of justice are not silent in this regard and explicitly endorse a constructivist view of normative justification as in (1). This is the case of James 2005 and Mackled-Garcia 2008.

intelligible, Cohen's critique of the concept of impure justice is not a critique of constructivism. To show that there is no need to engage with Cohen's critique itself. Such a critique has as a target a particular understanding of the concept of justice (i.e., (2)), an understanding that is not necessarily linked to the metaethical or normative methodology of constructivism (i.e., (1)).

Let me now turn to Cohen's critique of the concept of impure justice itself. As we have seen, the critique claims that fundamental principles of justice are to meet two conditions: the *fact-insensitivity condition* and the *single-value condition*. Since the principles that result from the original position do not meet these two conditions, Cohen argues, they are not fundamental principles of justice. I will make one observation about each condition.

Remember what the *fact-insensitivity condition* states: fundamental principles of justice are not justified on the basis that certain empirical facts are the case. This claim is supposed to threaten Rawls's justice-as-fairness and other practice-dependent conceptions of justice, as these conceptions assign a serious role to considerations of empirical fact in the justification of their principles. But this threat is merely apparent. For these conceptions can indeed presuppose further fact-insensitive principles without conceding anything substantial regarding the nature of justice.

These conceptions typically hold that a given principle *P* is adequate to regulate a practice partly because certain facts *F* of the practice are the case. Practice-dependent conceptions can meet the *fact-insensitivity condition*

by stating that this latter claim commits them to the following fact-insensitive principle:

(i) Nec, if F is the case, then P

This principle fulfils the requirement established by the *fact-insensitivity condition*, for it can be affirmed and justified regardless of whether F obtains (see James 2011: 7). As an illustration, let us consider Rawls's claim (*TJ*: 109-112) that the application of the principles of justice is conditional on whether the facts that he groups under the idea of the circumstances of justice obtain. Using the necessity operator as in its deontic interpretation (a standard way of stating that something is obligatory), we can formulate this principle as follows:

(ii) It is obligatory that if the circumstances of justice arise in a society with a liberal political culture, then such a society be regulated by the principles of justice-as-fairness.

This further principle fulfils the *fact-insensitivity condition*. It can be affirmed or justified regardless of whether the circumstances of justice are the case. So, even when within Rawls's theory the principles of justice-as-fairness are affirmed partly on the basis of the importance of considering the circumstances of justice, such a framework gives rise to a commitment to (ii). Rawlsians and defenders of similar practice-dependent conceptions of justice (either domestic, global, or for other domains) can accept that their conceptions are ultimately based on fact-insensitive principles like (i) and (ii) without conceding anything substantial on the debate about the nature of justice. If any single normative conception of justice can be ultimately based

on fact-insensitive conditional principles of this sort, then the *fact-insensitivity condition* does not illuminate anything particularly relevant either in the debate of the nature of justice or at the level of normative justification or meta-ethics. Pogge 2008: 458-461, after stating a similar observation, argues that Cohen's victory is 'Pyrrhic': 'Why do I call this a Pyrrhic victory? Because Cohen's metaethical triumph comes without the announce rescue of justice' (see also Raz 2011 for a similar point).

So let me turn to Cohen's *single-value condition*. Remember what the condition states: fundamental principles of justice are justified by appealing exclusively to considerations dependent on the value of justice. It is quite surprising that Cohen endorses this condition, as he is perfectly clear that the considerations he offers as a way of justifying his own preferred fundamental principle of justice (i.e., the luck/choice principle) are, in the end, entirely dependent on the value of *fairness*:

My own animating conviction in political philosophy with respect to justice is a conviction about distributive justice in particular. It is that an unequal distribution whose inequality cannot be vindicated by some choice of fault or desert on the part of (some of) the relevant affected agents is **unfair**, and therefore, *pro tanto*, unjust, and nothing can remove that particular injustice. (Cohen 2008: 7. Italics in the original, bold added)

In this passage Cohen states that distributive justice requires compliance with the value of fairness. In light of this remark, there are two interpretative options that create what I call *Cohen's dilemma*. Either Cohen is collapsing the value of fairness with that of justice or he is not. If he is, then Cohen is offering a circular consideration as a way of justifying why his preferred fundamental principle of justice is correct: justice requires the alleviation of

all inequalities that are not the result of choice *because* not doing so would be unjust. If, on the other hand, Cohen is not collapsing the values of fairness and justice, then he is appealing to another value rather than justice alone in order to justify the fundamental principle of justice he prefers, and thus he violates the *single-value condition*.

One may think that Cohen has simply made a mistake in the previous quote —i.e., appealing to a further value rather than justice to defend his preferred fundamental principle of justice—, a mistake that can be easily removed from his defence of the luck/choice principle. However, Cohen repeatedly mentions the importance of the value of fairness in support of the luck egalitarian understanding of justice. In the introduction of his book, he says again: ‘The Rawlsian approach to justice denatures justice, since it cannot recognize that *if something is unfair, then it’s to that extent of unjust*’ (Cohen 2008: 7-8. Italics added). Similar remarks about why fundamental principles of justice must comply with the value of fairness are made in Cohen 2011b: 5, 2011a: 120, 124-146. In fact, in one of the most comprehensive studies on Cohen’s luck egalitarianism, Michael Otsuka (2010) defends precisely the idea that at the core of such a view there are two impulses, *both of them dependent on the value of fairness*: the alleviation of exploitation and of brute-luck. I have exemplified the dependence of the latter (alleviating brute-luck) on the value of fairness. Otsuka makes the task of most of his paper to prove the dependence of the former (alleviating

exploitation) on the commitment to such a value. Otsuka summarises this by stating his main claim:

I shall conclude that Cohen's account of distributive justice has a non-egalitarian as well as an egalitarian aspect. The non-egalitarian aspect is his impulse to extinguish the influence of exploitation on the distribution of benefits, whereas the egalitarian aspect is his impulse to extinguish the influence of brute luck. *Each impulse arises from an underlying commitment to fairness.* (Otsuka 2010: 218. Italics added)

All this evidence supports the claim that Cohen's defence of the luck/choice principle depends on the value of fairness. In light of this, how are we to understand Cohen's endorsement of the *single-value condition*?

There are two ways for Cohen to avoid *Cohen's dilemma*: either he abandons the *single-value condition* or he renounces the 'underlying commitment to fairness' from which the impulse to support his fundamental principle of justice arises. It seems to me that the first option is preferable. In fact, it seems that if we cannot use other values in order to define fundamental principles of justice, we cannot define principles of justice at all. In Rawls's case, he makes use of the normative understanding of society as a fair system of social cooperation between free and equal persons; he may also suppose further fact-insensitive conditional principles as (i) and (ii). As we saw, these are normative considerations for which Rawls offers no justification and on which his conception of justice depends, just as Cohen's justification of the luck/choice principle depends on the value of fairness without further defence either.

Summing up, in this section I have defended three points: first, that Cohen is mistaken in thinking that his critique of the concept of *impure justice*

is a critique of constructivism. Second, that Cohen's *fact insensitivity condition* is trivial. Lastly, that Cohen's *single-value condition* creates *Cohen's dilemma*.

5.2 A *Constitutive VS an Indicative Account of Correctness*

In this section I will defend that there is a substantial difference in advancing constructivism as a metaphysical view and as an epistemic view regarding the correctness of normative principles (that is, a difference in holding that a given standard of correctness is *constituted* by being the result of a specified procedure and holding that satisfying a given standard of correctness is *indicated* by being the result of such a procedure). I will also defend that there can be normative justifications that are metaethically neutral. Finally, I will argue that the epistemic understanding of constructivism just mentioned states a trivial claim about normative justification.

Constructivist accounts stating the metaphysical claim of the constitutive structure of correctness are straightforwardly committed to a particular metaethical view, a view that is incompatible with other explanations of what *it is* for a principle to be correct —such as the various forms of moral realism, expressivism, fictionalism, and error theories. Contrarily, constructivist accounts stating the epistemic claim about the indication of correctness are not committed to a particular metaethical view in this regard. Instead, these accounts hold an understanding of normative justification that may be compatible with the usual list of metaethical options.

Of course, the idea that a normative justification can be metaethically neutral (as the epistemic understanding of constructivism is characterised to hold) has been contested. Some of the most powerful and paradigmatic critiques to such an idea can be found in authors as different as Rorty 1982: xvi, Dworkin 1996: 87-139, and Blackburn 1998: 295. Conversely, an early defence of the independence of normative justification from metaethical questions can be found in authors as different as Mackie 1977: 16, Rawls (*CP*: 287-288), and Korsgaard 1996a: 363-398 (although her own work shifted later towards the opposite position).⁴²

We can try to show the logical independence of normative justification by presenting what appears to be an extreme case: we can imagine a view of normative justification that is merely based on the coherence of all judgements regarding a specific discourse—for instance, moral discourse. This view would state that what justifies taking a judgement as correct is the very and only fact that it coheres with the rest of the judgements made by the same agent within moral discourse. On this view, nothing is said regarding the constitutive structure of the correctness of moral discourse's judgements. Because of that, this coherentist view of normative justification is compatible with a robust form of moral realism about moral discourse (i.e., the view that the truth of moral judgements is constituted by whether they refer to a mind-independent order of values or facts—see Shaffer-Landau 2003: 13-18, Enoch 2007: 21-51 and FitzPatrick 2008: 159-207). It is also compatible with

⁴² For a recent defence of the independence of metaethics from views about normative justification see Hussain & Shah 2006: 267 and Bloomfield 2009.

an error theory about moral discourse (i.e., the view that all moral judgements are false since there is no mind-independent order of values or facts to which they refer —see Mackie 1977).

To see why this is so, it is only needed to show that an agent can hold *the same view about normative justification* in two discourses for which such an agent holds *different metaethical views*. For instance, it is perfectly plausible to imagine an agent holding a coherentist view of normative justification (as the one just outlined) for both moral and scientific discourse while, at the same time, endorsing *realism* in the scientific discourse and an *error theory* in the moral discourse: on the one hand, this agent would believe that what justifies taking as correct judgements in the two discourses is the very and only fact that they cohere with the rest of the judgements made in each respective discourse. On the other hand, this agent would also believe that, while the scientific judgements that are justified in such a way *also* refer to a mind-independent order of facts, the moral judgements justified in the very same way do not refer to a mind-independent order of facts. That there is nothing irrational or incoherent about an agent adopting these beliefs, it seems to me, proves the logical independence of normative justification from metaethical commitments.

In fact, a quite common interpretation of Rawls's reflective equilibrium also assumes the independence of normative justification. Rawls (*PL*: 95-96) himself states that such a method of justification is used by his *political constructivism*, by *Kantian constructivism*, and by that form of moral

realism that he calls *rational intuitionism*. For now, suffice it to say that the last two are defined as opposed metaethical views regarding the constitution and origin of moral values.⁴³ Despite this opposition, Rawls claims, both use the same method of normative justification.

Having defended the idea that there can be metaethically neutral normative justifications, let me come back to our discussion regarding the two understandings of constructivism. For the purpose of clarity, let me refer to the metaethical understanding as *constructivism ME* and to the normative justification understanding as *constructivism NJ*. Recall what defines them:

Constructivism ME: Satisfying a given standard of correctness for the principles and judgements of a given discourse is *constituted* by being the result of a specified procedure.

Constructivism NJ: Satisfying a given standard of correctness for the principles and judgements of a given discourse is *indicated* by being the result of a specified procedure.

Notice that *constructivism ME* is a much more robust claim than *constructivism NJ*. *Constructivism NJ* is only a claim regarding normative justification for the principles of a given discourse: the fact that a principle would be issued by a given procedure *indicates* its correctness and thus justifies it. In contrast, *constructivism ME* is both a claim about normative justification and an explanation of the constitutive structure of correctness for a given discourse: the fact that a principle would be issued by a given procedure not only *indicates* its correctness. It is also *constitutive* of the correctness of such a principle that it is issued by the procedure.

⁴³ In § 5.3 we will see that, *pace* Rawls (*PL*: 102-130), *political constructivism* is also a metaethical view in this regard.

Now, recall that Cohen 2008: 274 defines constructivism as (1) ‘the view that a principle gains its normative credentials through being the result of a sound selection procedure’ and (2) the view ‘that proceeds by putting and answering the question “What rules of governance are to be adopted for our common social life”?’ In § 5.1 we saw that a constructivist view need not necessarily try to answer the question posed by (2) —as a constructivist view of *pure principles* of justice is perfectly intelligible. Considering the previous remarks, we are now in a position to disambiguate the meaning of (1) and, in particular, of the phrase *gaining normative credentials*: it may either imply a commitment to *constructivism ME* or to *constructivism NJ*. In fact, Cohen himself seems to be aware of the existence of these two different understandings of constructivism, as it is clear from the following passage:

I should acknowledge, here, a distinction that is of the first importance in philosophically but that will have no bearing on my own proceedings. I mean the distinction between the view that what it *is* for a principle to be valid, is that it is the product of some favored constructivist procedure; and a view according to which the constructivist procedure merely makes this principle valid [...] The stated distinction is at the pinnacle of metaethics, a pinnacle that my discussion does not reach. (Cohen 2008: 275-276. Italics in the original)

This passage makes clear that Cohen himself was aware of the fact that he was simply avoiding a discussion with constructivism throughout his critique of the concept of *impure justice*. However, Cohen’s remarks here are helpful for clarifying that the issues concerning constructivism are ‘at the pinnacle of metaethics’. Instead of being a view of how to understand the concept of justice, constructivism can either be a metaethical view or a view about

normative justification. Neither of these two understandings is reached by Cohen's critique.

The reception of constructivism has always oscillated between the two ways of understanding it (for recent discussion see O'Neill 2003, James 2007, Korsgaard 2008, Galvin 2010). Moreover, sometimes both understandings are confusingly conflated into a putatively single position. Ronzoni & Valentini 2008, and Ronzoni 2010 neatly exemplify this. For these authors argue, on the one hand, that constructivism is a view about the normative justification of moral principles that is agnostic regarding the existence of mind-independent moral facts. Ronzoni holds:

Constructivists believe that the authoritative normative principles *can* be justified —through the appeal to certain constraints of reason, and consequently to specific constructivist procedures, rather than to an independent order of facts [...] the most plausible stand one can take regarding the existence of independent moral truths is, I contend, one of agnosticism. (Ronzoni 2010: 74, 78. Italics in the original)

This passage explicitly endorses *constructivism NJ*, for it says that constructivism is a normative justification strategy that is neutral regarding metaethical discussions. However, on the other hand, these authors also endorse the constitutivist, non-metaethically neutral claim that characterises *constructivism ME*: 'From the standpoint of constructivist theory, intersubjective justifiability simply *is* the only form of objectivity available' (Ronzoni 2010: 74). A mixed conception like this cannot be coherent. If we make the metaphysical claim regarding the constructivist nature of moral principles' objectivity (that is, of what *it is* for a principle to be correct), we cannot remain neutral or agnostic regarding the nature of moral principles'

objectivity. This is incompatible with *constructivism NJ* since, by definition, such a view is an epistemic claim neutral regarding the nature and structure of objectivity and correctness.

So we must keep the two understandings separate. Moreover, there is reason to say that only *constructivism ME* can be a distinctive constructivist account proper. Sharon Street advances this position:

For notice that any view in ethics can say that the results of reasoning according to a certain procedure are correct. What makes a view constructivist is its claim that the results of reasoning according to a certain procedure are correct *because they issue from that procedure* —that to be correct *just is* to issue from that procedure. In other words, what is distinctive about a constructivist view is that they understand correctness to be *constituted* by emerging from certain procedure, and not merely *coincident* with it. (Street 2008: 212. Italics in the original)

Street is right in stating that constructivism can only be taken as a full-fledged distinctive view when is understood in a metaethical way. This is due to the fact that *constructivism NJ* is problematic.

There is a common objection (which goes back as far as Plato's *Euthyphron problem*) that this view cannot overcome. *Constructivists NJ* views always proceed in three stages: first, they specify a given procedure; second, they give arguments to support that certain principles would be issued by that procedure; finally, they conclude that, because of the arguments provided in the second stage, the principles would indeed be issued by the procedure. Crudely stated, the objection is: what justificatory work is being done by the third stage? In other words, what justificatory work does it take to stipulate that certain principles would be the result of the procedure? It seems that what justifies the principles are *the arguments*

provided in the second stage to show that they would be issued by the procedure. These arguments, if based on true premises and valid, prove that the principles are justified. Thus, it seems that what *constructivists NJ* are really stating is something like the following: ‘because these principles are *justified* by correct arguments, they would be issued by the proposed procedure’. However, this claim is clearly irrelevant for the purposes of normative justification. *Constructivism NJ* seems to be compatible with any view in ethics, as Street holds in the passage quoted, precisely because it is a trivial claim about normative justification.

In fact, Rawls’s theory is a perfect example of this problem. If the original position is interpreted in *constructivist NJ* fashion, then Rawls’s theory holds that the two principles of justice are justified because they issue from the original position. Rawls then exclusively focuses on providing arguments to show why justice-as-fairness’s principles would be chosen. But notice that all the arguments to that effect can be provided and accepted without appealing to the original position at all. Take some paradigmatic cases for some of the principles: Rawls’s (*TJ*: 62-63) argument showing the normative arbitrariness of one’s gender, sexual preference, social class and upbringing, in favour of the principle of fair equality of opportunity; Rawls’s (*TJ*: 90-91) argument stating that the ideal of fraternity is better respected by the difference principle; Rawls’s (*TJ*: 259-263) argument defending that time-preference is unjust in favour of the principle of just savings. Someone who rejects the putative justificatory force of *constructivism NJ* can still fully

accept all these arguments without showing incoherence. The fact that this is possible suffices to show the independent justificatory force of such arguments from the original position.

Perhaps *constructivists NJ* would like to say that there is another way of understanding their view that differs from the three-staged structure I have mentioned, and which can avoid the objection stated. I doubt that there is an alternative way. This is so because there is only one way in which *constructivists NJ* can show (without begging the question against their opponents) that the principles *they* prefer, as opposed to the ones they do not prefer, would be issued by their procedure: by providing arguments in favour of their principles. Thus, we are back to the idea that all the justificatory work is done in the second stage—i.e., by the arguments themselves.

Because of this objection, *constructivism ME* has better prospects than *constructivism NJ*.⁴⁴ When assessing *constructivist ME* views, it is useful to keep in mind Street's 2008: 208-209 distinction of what she calls *restricted constructivism* and *metaethical constructivism*. However, although the distinction is useful, the terminology employed is misleading for, as Street 2008: 217-219 herself states, both subtypes of constructivism affirm a metaethical claim. The difference between them is, rather, that the first one is

⁴⁴ It may still be that *constructivism ME* is also prey to the same objection I made to *constructivism NJ*—aside from other important weaknesses (see Shafer-Landau 2003: 29-51 and Enoch 2009). Here I will not analyse this further. For the purpose of this chapter, it is only relevant to show the inadequacy of *constructivism NJ* as a distinctive form of normative justification, and thus to show the inadequacy of interpreting the original position in *constructivist NJ* fashion. Moreover, in the next section I will argue that we should also avoid interpreting the original position in *constructivist ME* fashion too, forcing a deflationary interpretation. Therefore, it is ultimately irrelevant to this chapter whether *constructivism ME* is a successful metaethical explanation or not.

an account attempting to explain what *it is* for a specific set of principles and normative judgements to be correct —such as those of distributive justice (e.g., *CP*: 304) or moral wrongness (e.g., Scanlon 1998). Contrarily, the other subtype of constructivism is an account attempting to explain what *it is* for all principles and normative judgments to be correct (e.g., Korsgaard 1996b). In this sense, a less misleading terminology would label the two subtypes simply as *restricted constructivism ME* and *unrestricted constructivism ME*.

These can be defined as follows:

Restricted constructivism ME: Satisfying a given standard of correctness for a specific subset of normative principles and judgements is *constituted* by their being the result of a specified procedure.

Unrestricted constructivism ME: Satisfying a given standard of correctness for all normative principles and judgements is *constituted* by their being the result of a specified procedure.

Notice that while *restricted constructivism ME* is a metaethical explanation of the correctness of a specified set of principles and normative judgements, it remains silent regarding the metaethical status of the grounding normative principles used in its procedure of construction. So, for instance, if Rawls's justice-as-fairness is interpreted in *restricted constructivism ME* fashion, the result of the original position is thought of as constituting what it is for *impure principles* of justice to be correct, even though (as we saw § 5.1) it is not specified what is *constitutive* of the correctness of the grounding normative principles and judgements used in the original position — paradigmatically, of the normative idea of society understood as a fair system

of social cooperation between free and equal persons. Taking justice-as-fairness as a form of *restricted constructivism ME* makes it unproblematic to assume the correctness of such an idea since, in this interpretation, justice-as-fairness pretends only to be a metaethical explanation of the correctness of *impure principles* of justice, and not a metaethical explanation of *all* normative principles and judgements, including the one stating that a liberal society must be understood as a system of social cooperation between free and equal persons. A metaethical explanation for *all* normative principles and judgements (i.e. *unrestricted constructivism ME*) would resemble more what Rawls (*TJ*: 15, 95-96) calls ‘rightness as fairness’. However, Rawls is clear regarding his lack of engagement with such a project. Rather, this project is closer to Korsgaard’s 1996b: 122 explicit aims.⁴⁵

Summing up, in this section I have defended three points: first, that there are metaethically neutral normative justifications. Second, that we must distinguish *constructivism ME* (be it *restricted* or *unrestricted*) from *constructivism NJ*. Finally, that *constructivism NJ* states a superfluous claim about normative justification.

⁴⁵ More recently, such a project has been the concern of authors defending what it is now tellingly called *constitutivist* views in metaethics. These views pretend to prove the existence of normativity (i.e., that certain normative predicates are true) by appealing to certain aspects of what it is constitutive of being an agent. For instances and defences of constitutivist views in metaethics see Korsgaard 2009, Velleman 2009: 135-146. For recent attacks see FitzPatrick 2005, Hussain & Shah 2006, and Enoch 2006.

5.3 A Comprehensive VS a Political Conception of Justice

For the late Rawls (i.e. *PL*), the relevant standard of correctness is not *truth* but *reasonableness*. If the principles of justice-as-fairness prove to be reasonable, they are to be considered the *correct* principles for regulating the basic institutions of a liberal society.⁴⁶ Importantly, even if the standard of correctness for a position is *reasonableness* rather than *truth*, such a position can still hold a constructivist view regarding what *it is* for a principle to be reasonable —i.e., what is *constitutive* of a principle's reasonableness (as opposed to what merely *indicates* that a principle is reasonable).

As Street 2008: 212 states, it is an open interpretative question whether the late Rawls is claiming that reasonableness is *constituted* or merely *indicated* by being the result of the original position:

This raises a thorny interpretative question, with textual evidence on both sides. But the important point for our purposes is this. If Rawls is not making the constitutive claim, then justice as fairness does not qualify as a constructivist view [...] Regardless of whether Rawls is making the constitutive claim in *Political Liberalism*, it is clear that he is making it in his earlier 'Kantian Constructivism in Moral Theory', so his view there still stands as an example of restricted constructivism. (Street 2008: 212)

Street is referring to Rawls's clear *constructivist ME* passages such as the following: 'What distinguishes Kantian constructivism is essentially this: it specifies a particular conception of the person as an element in a reasonable procedure of construction, the outcome of which *determines* the content of

⁴⁶ This idea has been contested. See Raz 1990 and Cohen 2009 for an attack on the dispensability of truth in the justification of a conception of justice and in the domain of public reason. See Estlund 1998 and Quong 2011 for a defence. For the purposes of this chapter, it is irrelevant whether Rawls is mistaken on his claim that the adequate standard of correctness for a conception of justice must be *reasonableness* and not *truth*. In both cases, as we will see, a constructivist account can be provided.

first principles of justice' (CP: 304. Italics added). However, the final statement of his theory oscillates between the *constructivist ME* and the *constructivist NJ* interpretation. In fact, as we will now see, such an oscillation creates a tension in Rawls's late work that needs to be resolved. I will argue, *pace* Street 2008, that the elements of Rawls's late work pulling towards the *constructivist ME* interpretation must be excluded from justice-as-fairness, since the *constructivist NJ* interpretation suits much better the purposes of political liberalism. This is so because, according to political liberalism, it is of first importance to avoid metaethical controversies. Crudely stated, a view holding that being the result of the original position is *constitutive* of the correct principles of justice (i.e., constitutive of having the property of reasonableness) is a comprehensive view, whereas the view holding that being the result of the original position is an *indication* of being correct (i.e. an indication of having the property of reasonableness) is not. Therefore, so long as justice-as-fairness pretends to be a political rather than a comprehensive conception of justice, the *constructivist NJ* interpretation is preferable.

Having said that, remember that in § 5.2 I presented an objection against the claim that *constructivism NJ* is a successful view of normative justification. Accordingly, I will ultimately propose that justice-as-fairness should not be interpreted either as a form of *constructivism ME* or of *constructivism NJ*, but, rather, in deflationary fashion. First, let me contrast

some of the clearly incompatible elements associated with each of those two interpretations.

Let us briefly remember that *Political Liberalism* has two main aims, each of them being sensitive to a particular *fact* about liberal societies. First, (as I have already mentioned) sensitive to the *fact* of reasonable pluralism of comprehensive doctrines (i.e., pluralism of doctrines that disagree on their metaphysical views regarding the nature, structure and source of the realm of value), political liberalism aims at presenting justice-as-fairness as a political conception of justice. Rawls (*PL*: 11-15) states that such a conception must fulfil three characteristics: (i) it must regulate only the political domain; (ii) it must be presented as freestanding from comprehensive views; and (iii) it must be justified on ideas acceptable to all reasonable comprehensive doctrines, that is, ideas already present in the shared political culture of liberal societies. Second, sensitive to the *fact* of the pluralism of political conceptions of justice in liberal societies (where justice-as-fairness is just one amongst many), political liberalism aims at presenting justice-as-fairness as *the most reasonable conception* of such a family (see *PL*: xlv-xlix, 95-96).

Now, perhaps two elements are the most significant in favour of the *constructivist NJ* interpretation. The first one is stated, precisely, by the very second aim of political liberalism. For, on the one hand, Rawls claims that by being the result of the original position the two principles of justice are reasonable and thus correct. On the other hand, Rawls also claims that *other*

conceptions of justice, which by definition are not the result of the original position, are reasonable as well. This shows that being the result of the original position is not *constitutive* of having the property of reasonableness, rather, a mere *indication* of having such a property.

The second element favouring the *constructivist NJ* interpretation is Rawls's (*PL*) idea that his *political constructivism* is different to Kantian constructivism and compatible with rational intuitionism. Rawls's ideas regarding the latter are summarised in the following passage: 'Justice-as-fairness does not deny what [rational intuitionists] want to assert: namely, that the order of values displayed by constructivism [by the original position] is backed up by an independent order of values that *constitutes* itself' (*PL*: 95. *Italics added*). In turn, Rawls's ideas regarding Kantian constructivism are summarised in the following passage:

Another and deeper meaning of autonomy says that the order of moral and political values must be made, or itself be *constituted* by the principles and conceptions of practical reason. Let us refer to this as *constituted* autonomy. In contrast with rational intuitionism, *constituted* autonomy says that the so-called independent order of values does not constitute itself but is *constituted* by the activity, actual or ideal, of (human) practical reason. I believe this, or something like it, is Kant's view [...] Political liberalism must, of course, reject Kant's *constitutive* autonomy. (*PL*: 99-100. *Italics added*)

This second element on its own, it seems to me, rules out the *constructivist ME* interpretation; it seems sufficient for showing Rawls's intention of presenting justice-as-fairness as a *constructivist NJ* view compatible with several and opposite metaethical explanations regarding the origin, structure, and source of the realm of value. This is also consistent with Rawls's further and recurrent idea holding that the original position is the

most suitable form of 'public justification on questions of political justice given the fact of reasonable pluralism' (PL: 100).

However, there are well-developed elements in Rawls's final understanding of justice-as-fairness pulling towards exactly the opposite interpretation. Perhaps two of them are the most significant: the mirroring ideas of *full autonomy* (see PL: 77-81) and *political constructivism* (see PL: 89-125). I will show how these ideas (against Rawls's very intention) make justice-as-fairness comprehensive and thus inadequate for fulfilling the first aim of political liberalism mentioned above.

Leif Wenar (see Wenar 1995: 52-57) has already defended a similar view regarding *political constructivism*. Wenar claims that the constructivist account of *objectivity* is not compatible with the plurality of metaethical views in this regard, so that by endorsing a 'divisive account of the source of objective reasons, Rawls has recreated the problems of pluralism at a higher level' (Wenar 1995: 57).⁴⁷ I completely agree with this diagnosis. However, Wenar does not mention that the idea of *full autonomy* mirrors exactly that of the objectivity proper to *political constructivism*. Therefore, he misses an

⁴⁷ In fact, Wenar points at three comprehensively Kantian elements of Rawls's political liberalism (1) the burdens of judgement, (2) the reasonable moral psychology, and (3) *political constructivism*. These elements are comprehensively Kantian, Wenar holds, because at bottom they state that the *free exercise of practical reason* is, respectively, (1) the explanation of the fact of reasonable pluralism, (2) the source of motivation to follow political principles, and (3) the source of the normativity of political principles.

important point: that if *political constructivism* must go (so long as the political understanding of liberalism is to succeed), so must *full autonomy*.⁴⁸

The main purpose of *full autonomy* is to reject the previous interpretation of the original position as an extension of Kant's categorical imperative (see *TJ*: § 40). According to Kant (see *G*:4:441), an autonomous will is one that gives a law to itself; the autonomous will does that by acting on maxims that spring from and respect the different formulations of the categorical imperative. This seems to imply, at least according to the constructivist interpretation of Kant put forward by Rawls (*CP*: 303-358) and Korsgaard 1996b⁴⁹, that the autonomous will itself is the source of all moral values and principles or, alternatively, that being the outcome of correct practical reasoning is constitutive of all the moral values and principles there are. For clarity purposes, let me refer as *Kantian autonomy* to this form *unrestricted constructivism ME*.

Now, since *Kantian autonomy* is a comprehensive claim about the nature of value as a whole, it needs to be avoided if the political understanding of liberalism is to succeed. Rawls thinks this can be done by restricting his political liberalism to a narrow claim regarding the relevant scope of practical reasoning: a full politically autonomous person is taken to be the source of a subset of moral principles, namely, the source of principles

⁴⁸ In fact, Wenar does not mention the idea of *full autonomy* in his paper at all. This is also the case of Scheffler 1994, who makes very similar observations. Partly because of this, I centre my discussion on such an idea rather than on *political constructivism*.

⁴⁹ For a critique of such interpretation of Kant's moral philosophy see Wood 2008: 282-283. He now interprets Kant as defending a form of moral realism instead.

for regulating the institutions of society —the source, that is, of the principles of justice. The original position is thus a way of fleshing out what is constitutive of principles of justice for the political domain. Along with Kantian lines, the idea is that it is constitutive of such principles that they are the outcome of correct practical reasoning. Rawls says:

Note that it is not the parties but the citizens of a well-ordered society in their public life who are fully autonomous. This means that in their conduct citizens not only comply with the principles of justice, but *they also act from* this principles. Moreover, they recognize this principles as those that would be adopted in the original position [...] Thus, full autonomy is realized by citizens when *they act from principles of justice that specify the fair terms of social cooperation they would give to themselves* when fairly represented as free and equals. (PL: 77. Italics added)

Notice how well this description of citizens' *full autonomy* mirrors *Kantian autonomy*. According to Kant, an action that expresses moral value needs both to comply with moral requirements and be done because of those requirements (that is, *from* what Kant calls the *motive of duty* —see G:4:399). Likewise, as we have seen, Kant claims that an autonomous will is the one that gives a law to itself which, in turn, requires of such a will that it follow maxims springing from correct practical reasoning (i.e., the different formulations of the categorical imperative). In this passage, Rawls follows closely these characteristics of *Kantian autonomy*. The only modification made by Rawls is that such characteristics are restricted to the domain of the political, so that an agent is taken to be fully autonomous because she is the source of the principles regulating that domain and only when she acts from those principles. Contrary to what *Kantian autonomy* holds, a person can be

fully autonomous even if she does not comply with principles springing from correct practical reason in all other aspects of her moral life.

However, the idea of *full autonomy* creates problems for the success of the political understanding of liberalism. For one thing, the thesis that practical reason is the source of political principles is a straightforward metaethical claim. To see this, let me set into use the terminology deployed in the previous section. As we saw, *constructivist ME* views can be *restricted* or *unrestricted*. The former states what it is constitutive of a subset of principles and judgements' correctness (e.g., principles of distributive justice or of what we owe to each other). The latter states what is *constitutive* of all moral principles and judgements' correctness. However, both *restricted* and *unrestricted* forms of constructivism are of a same metaethical kind. They propose the same explanation for principles' correctness that is incompatible with those proposed by other metaethical theories such as the distinctive forms of moral realism, expressivism, fictionalism, and error theories. It is doubtful that a moral realist or an error theorist would accept that the will is the source of the principles regulating the political domain, i.e., that being the outcome of correct practical reasoning is constitutive of such principles. All such theorists believe in different and incompatible explanations in this regard (as long as political principles are a subset of moral principles more generally), all of them just as comprehensive as the one offered by Rawls's *restricted constructivism ME* recreated by the idea of *full autonomy*.

If it is indeed the case that *full autonomy* provides a comprehensive explanation of the source of political principles, then there is a strong reason to abandon it if the political understanding of liberalism is to succeed.⁵⁰ The easiest way of doing this is by interpreting the original position in *constructivist NJ* fashion —that is, as a way of offering a normative justification to the principles that are to regulate the political domain. However, as we have seen in the previous section, this interpretation is also problematic for, as I have argued, *constructivists NJ* views seem to be stating a superfluous claim about normative justification: ‘because these principles are *justified* by correct arguments, they would be issued the proposed procedure’. In light of this, I think we should opt for a third way of

⁵⁰ Let me distinguish my critique of *full autonomy* from a very usual critique of Rawls’s ideas on autonomy in light of his political liberalism. Many erred criticisms of Rawls’s political liberalism are based on the incorrect interpretative claim that, due to the fact of reasonable pluralism, Rawls argues that justice-as-fairness must renounce the importance of the value of autonomy. Kymlicka 1992 is the most prominent example of this incorrect interpretative claim. Such a claim rests on equivocation with regard to the term *autonomy*. According to Kymlicka 1989: 12-13, autonomy can be roughly understood as citizens’ capacity to choose amongst different conceptions of a good way of life. Autonomy thus stresses the importance that liberals place on the idea that a person’s life is of value only when lived ‘from the inside’. A very similar understanding of autonomy can be found in Raz’s 1986: 369 remarks: ‘The ruling idea behind the ideal of personal autonomy is that people should make their own lives. The autonomous person is (part) author of his own life’. In similar fashion, Wall 1998: 128 states that persons’ autonomy involves ‘making something out of their lives according to their own understanding of what it is valuable and worth doing’. Following the lead of these authors typically opposed to the political understanding of liberalism, let me assume the simple understanding of autonomy as the exercise of the capacity to choose amongst and reassess different ways of life. Call this understanding of the value of autonomy *liberal autonomy*. There is nothing in Rawls’s political liberalism arguing against *liberal autonomy*. In fact, many fundamental features of his conception of justice support it. The most relevant feature in this respect is Rawls’s (*PL*: 19) definition of persons’ second moral power: the capacity to pursue, revise and abandon a conception of the good. In Rawls’s liberal theory, thus, *liberal autonomy* is fleshed out as the second moral power of persons. So long as the institutions of the basic structure of society are concerned with the distribution of primary goods, and these goods are preconditions for the development of the two moral powers, Rawls’s theory endorses the commitment that the State should promote *liberal autonomy* amongst its citizens. Kymlicka 1995: 158 also mistakenly holds that Rawls’s political liberalism is an attempt to respond to putative communitarian critiques. Textual evidence favours the opposite claim (see *PL*: xxx, xxxii).

interpreting the original position that is not constructivist in either of the two senses. Call it *the deflationary interpretation*. Such view can be defined as follows:

Deflationary interpretation: the original position is a heuristic device whose function is to order all relevant reasons and arguments for justifying the principles that are to regulate the institutions of a liberal society.

Notice that, by taking the original position merely as a heuristic device for the exposition all relevant liberal reasons in favour of principles of justice, the *deflationary interpretation* is consistent with the claim that those reasons can also be presented without appealing to such a procedure. Therefore, this interpretation is based on the idea that all the justificatory work of the principles is done by the liberal reasons and arguments themselves.

In fact, this third understanding of the original position is not new. Rawls himself sometimes speaks in favour of it, as when he says that the original position is merely a *model* of the reasons we, here and now, think are relevant for the justification of liberal principles (see *JF*: 17, *PL*: 25). This idea implies that, while the model is dispensable, the reasons that it models are not. Therefore, in opposition to its *constructivist ME* rival, the *deflationary interpretation* holds that the original position is not an explanation of the constitutive structure of principles of justice's correctness or objectivity. In turn, in opposition to its *constructivist NJ* rival, the *deflationary interpretation* holds that nothing is added to the normative justification of the principles by the fact that those principles would be *agreed* upon the original position.

Perhaps *constructivists NJ* would like to push a more robust claim of justification than the one made by the *deflationary interpretation*, while still trying to remain metaethically neutral. They could say that the *only* correct understanding of justification —given the fact of reasonable pluralism (*PL*: 99) or given the fact of disagreement in metaethics (Ronzoni 2010: 79)— *is* the constructivist understanding. This is tantamount to claiming that the only way of *justifying* principles of justice is by appealing to the original position —regardless of whether it is constitutive of them to be the result of such a procedure. If this claim is correct, we could not interpret the original position in a *deflationary* fashion.

However I find this claim puzzling for three reasons. First, it flies in the face of the objection I presented in § 5.2. As we saw, we can accept all the arguments in favour of the principles of justice-as-fairness *and* reject *Constructivism NJ* without showing incoherence. Second, stating that there is only one way of justifying principles, and that such a way is constructivist in nature, seems tantamount to stating that it is in the nature of the objectivity of those very principles to be issued by a procedure of construction. But, as we saw in § 5.2 regarding Ronzoni and Valentini’s mixed conception, this claim about objectivity is straightforwardly metaethical and thus incompatible with the other usual explanations of principles’ objectivity. Finally, stating that there is only one way of justifying principles seems to go against the spirit of the very idea of an overlapping consensus about such principles (*PL*: 131-168). According to such an idea, there can be many

different moral reasons in support of the principles of justice. It sounds quite implausible to interpret this idea as stating that there are many moral reasons to accept the *one and only* correct justification of the principles of justice. Rather, a more plausible interpretation of the idea of an overlapping consensus is that there are *many* moral justifications to endorse and comply with the principles of justice —so long as these are based on moral rather than prudential reasons. Only in this way, it seems to me, justice-as-fairness ‘can win its support by addressing each citizen’s reason, as explained within its own framework’ (PL: 143), as the idea of an overlapping consensus states.⁵¹

Summing up, in this section I have defended two claims. First, that because of the two aims of political liberalism (i.e., that justice-as-fairness is a political conception of justice and the most reasonable one within such a family of conceptions), the *constructivist NJ* interpretation of the original position is preferable to its *constructivist ME* rival (and thus we should abandon Rawls’s idea of *full autonomy*). Second, that because of the internal problems of any *constructivist NJ* position, the *deflationary interpretation* of the original position is preferable to its *constructivist NJ* rival.

⁵¹ Something similar to this point is defended by Scheffler 1994: 15-22. He conclusively claims: ‘if an overlapping consensus on liberal principles can be achieved in modern democracies, then accepting any of the doctrines included in such a consensus must give one reason to support a liberal scheme. In other words, the distinctive contribution of political liberalism may be to suggest that there are many ways to arrive to at liberal principles and that that very fact is a source of liberalism’s strength.’ (Scheffler 1994: 22)

5.4 Conclusive Remarks

This concludes my assessment and defence of Rawls's practice-dependent conception of domestic justice. As we have seen, while such a conception can overcome Cohen's recent criticisms, it must nevertheless refrain from a commitment to a particularly robust understanding of constructivism either at the meta-ethical level or at the level of normative justification; otherwise the project of offering a political conception of justice for the regulation of liberal institutions would not succeed.

References

- Abdel-Nour, F. 2003. "National Responsibility". *Political Theory*, 31/5, pp. 693-719.
- Abizadeh, A. 2004. "Historical Truth, National Myths, and Liberal Democracy: On the Coherence of Liberal Nationalism". *The Journal of Political Philosophy*, 12/3, pp. 291-313.
- Ackerman, B. 1997. "Temporal Horizons of Justice". *The Journal of Philosophy*, 94, pp. 299-317.
- Acosta, M. & Ennelin, E. 2006. "'The Mexican Solution' to Transitional Justice". In Roth-Arriaza, N. & Mariezcurrena, J., *Transitional Justice in the Twenty-First Century. Beyond Truth and Justice*. Cambridge University Press, pp. 94-119.
- Agger, I. & Jensen, S. 1996. *Trauma and Healing under State Terrorism*. Zed books.
- Anderson, E. 1999. "What Is the Point of Equality". *Ethics*, 109/2, pp. 287-337.
- _____. 2007. "How Should Egalitarians Cope with Market Risk?" *Theoretical Enquiries in Law*, 9, pp. 239-270.
- _____. 2008. "Expanding the Egalitarian Toolbox: Equality and Bureaucracy". *Proceedings of the Aristotelian Society, Supplementary Volume*, 82, pp. 139-160.
- Arendt, H. 2006. *Eichmann in Jerusalem. A Report on the Banality of Evil*. Penguin Books.

- Arneson, R. 1999. "Against Rawlsian Equality of Opportunity". *Philosophical Studies*, 93/1, pp. 77-112.
- Backer, D. et al. 1995. "Therapy with the Victims of Political Repression in Chile: The Challenges of Social Reparations". In Kritz, J. *Transitional Justice. How Emerging Democracies Reckon with Former Regimes*. US Institute of Peace.
- Baldwin, T. 2008. "Rawls and Moral Psychology". In Shafer-Landau, R., *Oxford Studies in Metaethics, Vol. 3*. Oxford University Press, pp. 247-271.
- Barry, B. 2001. *Culture and Equality. An Egalitarian Critique to Multiculturalism*, Polity Press.
- Ben-Ze'ev, A. 2002. "Are Envy, Anger, and Resentment Moral Emotions?" *Philosophical Explorations*, 5/2, pp. 148-154.
- Bird, C. 2008. "Self-respect and Respect for Others". *European Journal of Philosophy*, 18/1, pp. 17-40.
- Blackburn, S. 1998. *Ruling Passions*. Clarendon Press.
- Bloomfield, P. 2009. "Archimedeanism and Why Metaethics Matters". *Oxford Studies in Metaethics, Vol. 4*, pp. 283-302.
- Blustein, J. 2008. *The Moral Demands of Memory*. Cambridge University Press.
- Boettcher, J. 2009. "Race, Ideology, and Ideal Theory." *Metaphilosophy*, 40/2, pp. 337-259.
- Borneman, J. 2005. "Public Apologies as Performative Redress". *SAIS review of international affairs* 25(2), pp. 53-65.
- Buss, S. 1999. "Respect for Persons". *Canadian Journal of Philosophy*, 4, pp. 517-550.
- Boxill, B. 1976. "Self-Respect and Protest". *Philosophy and Public Affairs*, 6/1, pp. 58-69.
- _____. 2003. "A Lockean Argument for Reparations". *Journal of Ethics*, 7, pp. 63-91.

- Buchanan, A. 2004. *Justice, Legitimacy, and Self-Determination*. Oxford University Press.
- Butt, D. 2009. *Rectifying International Injustice. Principles of Compensation and Restitutions between Nations*. Cambridge University Press.
- Campbell, J. 1994. *Past, Space, and Self*. MIT Press.
- Chapman, A. R. & Ball, P. 2001. "The Truth of Truth Commissions: Comparative Lessons from Haiti, South Africa, and Guatemala," *Human Rights Quarterly* 23/1, pp. 1-43.
- Cohen, A. 2009. "Compensation for Historic Injustices. Completing the Boxill and Sher Argument." *Philosophy and Public Affairs*, 37/1, pp. 81-102.
- Cohen, G. A. 1992. "Incentives, Inequality, and Community". In *The Tanner Lectures on Human Values, Vol. 13*. Salt Lake City: University of Utah Press, pp. 261-329.
- _____. 1997. "Where the Action Is: on the Site of Distributive Justice". *Philosophy and Public Affairs*, 26, pp. 3-30.
- _____. 2008. *Recovering Justice and Equality*. Harvard University Press.
- _____. 2011a. "Fairness and Legitimacy in Justice, and: Does Option Ever Preserve Justice?" In *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*. Princeton University Press, pp. 3-43.
- _____. 2011b. "On the Currency of Egalitarian Justice". In *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*. Princeton University Press, pp. 116-123.
- Cohen, J. 2009. "Truth and Public Reason". *Philosophy and Public Affairs*, 37, pp. 1-42.
- Constant, R. & Foxx, R. 2003. "Self-Respect. A Neglected Concept". *Philosophical Psychology*, 16/2, pp. 247-288.
- Conway, M. & Pleydell-Pearce. 2000. "The Construction of Autobiographical Memories in the Self-Memory System. *Psychological Review*, 107/2, pp. 261-288.
- Daniels, N. 2003. "Democratic Equality: Rawls's Complex Egalitarianism". In *The Cambridge Companion to Rawls*, pp. 241-276.

- Danto, A. 1985. *Narrative and Knowledge*, Columbia University Press.
- Darwall, S. 1977. "Two Kinds of Respect". *Ethics*, 88/1, pp. 136-149.
- de Grieff, P. 2006a. "Introduction: Repairing the Past: Compensation for Victims of Human Rights". In de Grieff, P. (ed), *The Handbook of Reparation*, Oxford University Press, pp. 1-15.
- _____. 2006b. "Justice and Reparations". In de Grieff, P. (ed.), *The Handbook of Reparation*, Oxford University Press, pp. 451-477.
- Dennet, D. 1989. "The Origins of the Self". *Cogito* 3, pp. 163-173.
- _____. 1992. "The Self as a Center of Narrative Gravity". In Kessel, F, et al (Ed.) *Self and Consciousness. Multiple Perspectives*. Earlbaum Associates, pp. 103-115.
- Dillon, R. 1992. "How to Lose Self-Respect". *American Philosophical Quarterly*, 29/2, pp. 125-139.
- Dworkin, R. 1996. "Objectivity and Truth: You Better Believe It". *Philosophy and Public Affairs*, 25/2, pp. 87-139.
- _____. 2000. *Sovereign Virtue. The Theory and Practice of Equality*. Harvard University Press.
- Elster, J. 2004. *Closing the Books. Transitional Theory in Historical Perspective*. Cambridge University Press.
- Enoch, D. 2006. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Agency". *Philosophical Review*, 115, pp. 169-198.
- _____. 2007. "An Outline of an Argument for Robust Metanormative Realism". *Oxford Studies in Metaethics, Vol. 2*. Oxford University Press.
- _____. 2009. "Can There Be a Global, Interesting, Coherent Constructivism about Practical Reason?". *Philosophical Explorations*, 12/3, pp. 319-339.
- English, J. 1977. "Justice between Generations". *Philosophical Studies*, 31, pp. 91-104.

- Estlund, D. 1998. "The Insularity of the Reasonable. Why Political Liberalism Should Accept the Truth". *Ethics*, 108, pp. 252-75.
- Eyal, N. 2005. "'Perhaps the Most Important Primary Good': Self-Respect and Rawls's Principles of Justice". *Politics, Philosophy, and Economics*, 4/2, pp. 195-219.
- Farrelly, C. 2007. "Justice in Ideal Theory: a Refutation". *Political Studies*, 55, pp. 844-64.
- FitzPatrick, W. 2005. "The Practical Turn in Ethical Theory: Korgaard's Constructivism, Realism, and the Nature of Normativity". *Ethics*, 115, pp. 651-691.
- _____. 2008. "Robust Ethical Realism, Non-Naturalism, and Normativity". *Oxford Studies in Metaethics*, Vol. 3. Oxford University Press.
- Fivush, R. & Nelson, C. 2004. "The Emergence of Autobiographical Memory: a Socio Cultural Developmental Theory". *The Psychological Review*, 111/2, pp. 486-511.
- Frankfurt, H. 1999. *Necessity, Volition, and Love*. Cambridge University Press.
- Freeman, S. 2007a. *Rawls*. Routledge.
- _____. 2007b. *Justice and the Social Contract*. Oxford University Press.
- Fuller, L. 2011. "Burdened Societies and Transitional Justice". *Ethical Theory and Moral Practice*. Forthcoming.
- Galston, W. 1991. *Liberal Purposes. Goods, Virtues, and Duties in the Liberal State*. Cambridge University Press.
- Galvin, R. 2010. "Rounding Up the Usual Suspects: Varieties of Kantian Constructivism in Ethics". *The Philosophical Quarterly*, 61/242, pp. 16-36.
- Gillespie, C. 1991. *Negotiating Democracy. Politicians and Generals in Uruguay*. Cambridge University Press.
- Grandin, G. & Miller, T. 2007. "Truth Commissions: State Terror, History and Memory." *Radical History Review*, 93.

- Gueda, V. 2010. 'The Old Colonialism Ends, the New Colonialism Begins'. In Beezley, W; Meyer, M. (eds.) *The Oxford History of Mexico*. Oxford University Press, pp. 265-284.
- Guembe, J. M. 2006. "Economic Reparations for Grave Human Rights Violations: The Argentina Experience". In de Grief, P. (ed.), *The Handbook of Reparation*, Oxford University Press, pp. 21-45.
- Hamber, B. 2006. "Narrowing the Micro and the Macro: A Psychological Perspective on Reparations in Societies in Transition. In de Grief, P., *The Handbook of Reparation*, Oxford University Press, pp. 560-588.
- _____. 2010. *Transforming Societies after Political Violence. Truth, Reconciliation, and Mental Health*. Springer.
- Herman, B. 1993. *The Practice of Moral Judgement*. Harvard University Press.
- _____. 2007. *Moral Literacy*. Harvard University Press.
- Herman, J. 1992. *Trauma and Recovery. The Aftermath of Violence —from Domestic Violence to Political Terror*. Basic Books.
- Herstein, O. 2008. "Historic Injustice and the Non-Identity Problem". *Law and Philosophy*, 27, pp. 505-531.
- Herzog, D. 2006. *Cunning*. Princeton University Press.
- Hill, T. 1973. "Servility and Self-Respect". *The Monist*, 57/1, pp. 87-104.
- _____. 1985. "Self-Respect Reconsidered". Reprinted in Dillon, R (ed). *Dignity, Character, and Self-Respect*. Routledge, 1995, pp. 76-93.
- Hinton, T. 2001. "Must Egalitarians Choose Between Fairness and Respect?". *Philosophy and Public Affairs*, 30/1, pp. 72-87.
- Holtermann, J. 2010. "The End of 'the End of Impunity'? The International Criminal Court and the Challenge From Truth Commissions." *Res Publica* 16/2, pp. 209-225.
- Hussain, N & Shah, N. 2006. "Misunderstanding Metaethics: Korsgaard's Rejection of Realism." *Oxford Studies of Metaethics*, Vol. 1, pp. 265-294.

- Jelin, E. 2007. "Public Memorialization in Perspective: Truth, Justice, and Memory of the Past in the Southern Cone of South America". *The International Journal of Transitional Justice*. 1, pp. 138-156.
- James, A. 2005. "Constructing Justice from Existing Practice. Rawls and the Status-Quo." *Philosophy and Public Affairs*, 33/3, pp. 281-316.
- _____. 2007. "Constructivism about Practical Reason". *Philosophy and Phenomenological Research*, 74/2, pp. 302-325.
- _____. 2011. "Deflating Fact Insensitivity". Unpublished.
- Jones, A. 2004. "Gendercide and Genocide". In Jones, A., (ed), *Gendercide and Genocide*, Vanderbilt University Press, pp. 1-37.
- Kant, I. 1785. *Groundwork of the Metaphysics of Morals*. Trans. by Mary Gregor. Cambridge University Press, 1996.
- Korsgaard, C. 1996a. *Creating the Kingdom of Ends*. Cambridge University Press.
- _____. 1996b. *The Sources of Normativity*. Cambridge University Press.
- _____. 2008. "Realism and Constructivism in the Twentieth-Century Moral Philosophy." In *The Constitution of Agency*. Oxford University Press, pp. 302-326.
- _____. 2009. *Self-Constitution. Agency, Identity and Integrity*. Oxford University Press.
- Kripke, S. 1980. *Naming and Necessity*. Harvard University Press.
- Kutz, C. 2004. "Justice in Reparations. The Cost of Memory and the Value of Talk." *Philosophy and Public Affairs*, 32/2, pp. 277-312.
- Kymlicka, W. 1989. *Liberalism, Community, and Culture*. Oxford University Press.
- _____. 1992. "Two Models of Pluralism and Tolerance" *Analyse und Kritik*, 14/1, pp. 33-56. Reprinted in Heyd, D. *Toleration*, Princeton University Press, 1996.
- _____. 1995. *Multicultural Citizenship*. Oxford University Press.

- _____. 2002. *Contemporary Political Philosophy*. Second Edition. Oxford University Press.
- _____. 2007. *Multicultural Odysseys. Navigating the New International Politics of Difference*. Oxford University Press.
- La Caze, M. 2001. "Envy and Resentment". *Philosophical Explorations*. 1, pp. 31-45.
- Lazenby, H. 2010. "One Kiss Too Many? Giving, Luck-Egalitarianism, and Other-Affecting Choice". *The Journal of Political Philosophy*, 18/3, pp. 271-286.
- Lawford-Smith, H. 2010. "Debate: Ideal Theory—A Reply to Valentini". *The Journal of Political Philosophy*, 18/3, pp. 357-368.
- Lira, E. 2006. "The Reparations Policy for Human Rights Violation in Chile". In de Grief, P. (ed.), *The Handbook of Reparation*, Oxford University Press, pp. 56-95.
- Lutz, E. 2007. "Transitional Justice: Lessons Learned and the Road Ahead". In Roht-Arriaza, N; Mariezcurrena, J. (Eds.). *Transitional Justice in the Twenty-First Century: Beyond Truth versus Justice*, pp. 325-341
- Lykes, B. & Mersky, M. 2006. "Reparations and Mental Health: Psychological Interventions Towards Healing, Human Agency, and Rethreading Social Realities". In de Grief, P. *The Handbook of Reparations*, Oxford University Press.
- MacIntyre, A. 1981. *After Virtue*. Duckworth.
- Mackie, J. 1977. *Ethics: Inventing Right and Wrong*. Penguin.
- Meckled-Garcia, S. 2011. "On the Very Idea of Cosmopolitan Justice: Constructivism and International Agency". *The Journal of Political Philosophy*.
- Margalit A. & Raz, J. 1995. "National Self-Determination". In Raz, J. *Ethics in the Public Domain*. Clarendon, pp. 125-146.
- Margalit, A. 1994. *The Decent Society*. Harvard University Press.
- _____. 2002. *The Ethics of Memory*. Harvard University Press.

- Martin-Baró, I. 1994. *Writings for a Liberation Psychology*. Harvard University Press.
- Massey, S. J. 1983. "Is Self-Respect a Moral or Psychological Concept?", *Ethics*, 93/2, pp. 246-261.
- Mason, A. 2010. "Rawlsian Theory and the Circumstances of Politics". *Political Theory*, 38/5, pp. 658-683.
- McGaught, J. 2003. *Memory and Emotion. The Making of Lasting Memories*. Columbia University Press.
- Meyers, D. 1989. *Self, Society and Personal Choice*. Columbia University Press.
- Milgram, S. 1974. *Obedience to Authority*. Harper.
- Miller, D. 1995. *On Nationality*. Oxford University Press.
- Mills, C. 2005. "Ideal Theory as Ideology". *Hypatia*, 20/3, pp. 165-184.
- Minow, M. 1998. *Between Vengeance and Forgiveness*. Beacon.
- _____. 2002. "Breaking Cycles of Hatred". In Minow, M. & Rosenblum, L. *Breaking Cycles of Hatred: Memory, Law, and Repair*. Princeton University Press.
- Moody-Adams, M. 1992. "Race, Class, and the Social Construction of Self-Respect". Reprinted in Dillon, R (ed). *Dignity, Character, and Self-Respect*. Routledge, 1995, pp. 271-289.
- Morris, C. 1984. "Existential Limits to the Rectification of Past Wrongs". *American Philosophical Quarterly*, 21/2, pp. 175-182.
- Munoz-Dardé, V. 2005. "Equality and Division: Values in Principle". *Proceedings of the Aristotelian Society*, 79, pp. 255-284.
- Neisser, U. 1994. "Self-Narratives: True and False". In Neisser, U & Fivush, R (eds). *The Remembering Self. Construction and Accuracy in Self-Narrative*. Cambridge University Press, pp. 1-19.
- Neimeyer G. J. & Metzler, A. E. 1994. "Personal Identity and Autobiographical recall. In Neisser, U & Fivush, R (eds). *The Remembering Self. Construction and Accuracy in Self-Narrative*. Cambridge University Press, pp. 105-136.

- O'Neill, O. 2003. "Constructivism in Rawls and Kant". *The Cambridge Companion to Rawls*, pp. 347-367.
- Otsuka, M. 2010. "Justice as Fairness: Luck-Egalitarian, Not Rawlsian". *The Journal of Ethics*, 14, pp. 217-230.
- Pauder-Studer, H & Velleman, D. 2011. "Distortions of Normativity". *Ethical Theory and Moral Practice*, 14, pp. 329-356.
- Parfit, D. 1984. *Reasons and Persons*. Oxford University Press.
- _____. 1997. "Equality and Priority". *Ratio*, 10/3, pp. 202-221.
- _____. 2011. *On What Matters*. Oxford University Press.
- Payne, L. 2008. *Unsettling Accounts. Neither Truth nor Reconciliation in Confessions of State Violence*. Duke University Press.
- Pasternak, A. 2011. 'Sharing the Cost of Political Injustice'. *Politics, Philosophy and Economics*, 10/2, pp. 188-210.
- Pogge, T. 2000. "The International Significance of Human Rights". *Journal of Ethics*, 4, pp. 45-69.
- _____. 2008. "Cohen to the Rescue!". *Ratio*, XXI/4, pp. 454-475.
- Quinn, J. & Freeman, M. 2003. "Lessons Learned: Practical Lessons Gleaned from inside the Truth Commissions of Guatemala and South Africa". *Human Rights Quarterly*, 25/4, pp. 1117-1149.
- Quong, J. 2011. *Liberalism without Perfection*. Oxford University Press.
- Rawls, J. 1999a. *A Theory of Justice*. Revised Edition. Harvard University Press.
- _____. 1999b. *The Law of Peoples*. Harvard University Press.
- _____. 1999c. *Collective Papers*. Harvard University Press.
- _____. 2001. *Justice as Fairness. A Restatement*. Harvard University Press.
- _____. 2005. *Political Liberalism*. Columbia University Press.
- Raz, J. 1986. *The Morality of Freedom*, Oxford University Press.

- _____. 1990. "Facing Diversity: The Case of Epistemic Absence". *Philosophy and Public Affairs*, 19/1, pp. 3-46.
- _____. 1995. "Multiculturalism: A Liberal Perspective". In Raz, J., *Ethics in the Public Domain*. Clarendon, pp. 170-192.
- _____. 2011. "Rescuing Jerry from Basic Principles". Unpublished.
- Robeyns, I. 2008. "Ideal Theory in Theory and Practice". *Social Theory and Practice*, 34, pp. 341-342.
- Robins, N. 2005. *Native Insurgencies and the Genocidal Impulse in the Americas*. Indiana University Press.
- Ronzoni, M. & Valentini, L. 2008. "On the Meta-Ethical Status of Constructivism: Reflections on G. A. Cohen's 'Facts and Principles'". *Politics, Philosophy and Economics*, 7/4, pp. 403-422.
- Ronzoni, M. 2010. "Constructivism about Practical Reason: on Intersubjectivity, Abstraction, and Judgement". *Journal of Moral Philosophy*, 7, pp. 74-104.
- Roper, S; Barria, L. 2009. "Why Do States Commission the Truth? Political Considerations in the Establishment of African Truth and Reconciliation Commissions," *Human rights review* 10/3, pp. 373-91.
- Rorty, R. 1982. *Consequences of Pragmatism*. University of Minnesota Press.
- Rosen, G. 2003. "Culpability and Ignorance". *Proceedings of the Aristotelian Society*, 103/1, pp. 61-84.
- _____. 2004. "Skepticism about Moral Responsibility". *Philosophical Perspectives*, 18/1, pp. 295-313.
- Sabini, J. Silver, M. 2005. "Lack of Character? Situationism Critiqued". *Ethics*, 115, pp. 535-562.
- Sach, D. 1981. "How to Distinguish Self-Respect from Self-Esteem". *Philosophy and Public Affairs*, 10/4, pp. 346-360.
- Sangiovanni, A. 2008. "Justice and the Priority of Politics to Morality". *Journal of Political Philosophy*, 36/2, pp. 137-164.
- Scanlon, T. 1998. *What We Owe to Each Other*. Harvard University Press.

- _____. 2003. *The Difficulty of Tolerance. Essays in Political Philosophy*. Cambridge University Press.
- Scheffler, S. 1994. "The Appeal of Political Liberalism". *Ethics*, 105/1, pp. 4-22.
- _____. 2003. "What Is Egalitarianism". *Philosophy and Public Affairs*, 31/1, pp. 5-39.
- _____. 2005a. "Choice, Circumstances, and the Value of Equality". *Politics, Philosophy and Economics*, 4/1, pp. 5-28.
- _____. 2005b. "The Division of Moral Labour". *Proceedings of the Aristotelian Society, Supplementary Volume*, 79/1, pp. 255-284.
- _____. 2006. "Is the Basic Structure Basic?" In Synowhich, C. (ed) *The Egalitarian Conscience. Essays in Honour of G. A. Cohen*, Oxford University Press, pp. 102-129.
- Schwartzman, L. 2006. *Challenging Liberalism: Feminism as Political Critique*. Penn State Press.
- Sen, A. 1977. "Rational Fools. A Critique of the Behavioral Foundations of Rational Choice Theory." *Philosophy and Public Affairs*, 6/4, 317-344.
- _____. 2006. "What Do We Want from a Theory of Justice?" *Journal of Philosophy*, 103, 215-238.
- _____. 2009. *The Idea of Justice*. Harvard University Press.
- Shaffer-Landau, R. 2003. *Moral Realism. A Defence*. Oxford University Press.
- Sher, G. 1975. "Justifying Reverse Discrimination in Employment". *Philosophy and Public Affairs*, 4, pp. 159-170.
- _____. 1977. "Groups and Justice". *Ethics*, 87, pp. 174-181.
- _____. 1979. "Compensation and Transworld Identity". *The Monist*, 62/3, pp. 379-391.
- _____. 1981. "Ancient Wrongs and Modern Rights". *Philosophy and Public Affairs*, 10/1, pp. 3-17.
- _____. 1997. *Approximate Justice. Studies in Non-Ideal Theory*. Roman & Littlefield.

- _____. 1999. "Diversity". *Philosophy and Public Affairs*, 28/2, pp. 85-104.
- _____. 2005. "Transgenerational Compensation". *Philosophy and Public Affairs*, 3/2, pp. 181-200.
- Shiffrin, S. 2004. "Race, Labour, and the Fair Equality of Opportunity Principle". *Fordham Law Review*, 72, pp. 1643-1675.
- _____. 2010. "Incentives, Motives, and Talents". *Philosophy and Public Affairs*, 38/2, pp. 111-142.
- Simmons, J. 1995. "Historical Rights and Fair Shares". *Law and Philosophy*, 14/2, pp. 149-184.
- _____. 2010. "Ideal and Non-Ideal Theory". *Philosophy and Public Affairs*, 38, pp. 5-36.
- Smyth, M. B. 2007. *Truth Recovery and Justice after Conflict. Managing Violent Past*. Routledge.
- Stemplowska, Z. 2008. "What's Ideal about Ideal Theory?". *Social Theory and Practice*, 34, pp. 319-340.
- Strawson, G. 2004. "Against Narrativity". *Ratio*, 17, 4, pp. 428-452.
- Street, S. 2008. "Constructivism about Reasons". *Oxford Studies in Metaethics*, Vol. 3, Oxford University Press, pp. 207-247.
- Svoboda, E. et al. 2006. "The Functional Neuroanatomy of Autobiographical Memory: a Meta-Analysis". *Neuropsychologia*, 44/12, pp. 2189-2208.
- Swift, A. 2008. "The Value of Philosophy in Non-Ideal Circumstances". *Social Theory and Practice*, 34, pp. 319-340.
- Tamir, Y. 1993. *Liberal Nationalism*. Princeton University Press.
- Tan, K. 2008. "A Defense of Luck Egalitarianism". *The Journal of Philosophy*, 105/11, pp. 665-690.
- Taylor, C. 1989. *Sources of the Self*. Cambridge University Press.
- Taylor, R. 2009. "Rawlsian Affirmative Action". *Ethics*, 119/3, pp. 476-506.
- Temkin, L. 1993. *Inequality*. Oxford University Press.

- Thomas, L. 1983. "Self-Respect: Theory and Practice". In Harris, T. (ed), *Philosophy Born Struggle: Anthology of Afro-American Philosophy from 1917*. Hunt Publishing House Company.
- Thompson, J. 2001. "Historical Injustice and Reparation: Justifying the Claims of Descendants". *Ethics*, 112, pp. 114-131.
- _____. 2006. "Collective Responsibility for Historical Injustice". *Midwest Studies in Philosophy*, 30, pp. 154-167.
- _____. 2009. *Intergenerational Justice*. Routledge.
- Titelbaum, M. 2008. "What Would a Rawlsian Ethos Look like?" *Philosophy and Public Affairs*, 36/3, pp. 289-322.
- Tomlin, P. 2008. "Envy, Facts and Justice: a Critique of the Treatment of Envy in Justice as Fairness". *Res Publica*, 14, pp. 101-116.
- Tomuschat, C. 2001. "Clarification Commission in Guatemala," *Human Rights Quarterly* 23/2, pp. 233-258.
- Valentini, L. 2009. "On the Apparent Paradox of Ideal Theory". *The Journal of Political Philosophy*, 17, pp. 332-55.
- _____. 2011. "Global Justice and Practice-Dependence: Conventionalism, Institutionalism, Functionalism". *The Journal of Political Philosophy*, pp. 399-418.
- Vandervort, B. 2006. *Indian Wars of Mexico, Canada and the United States, 1812-1900*. Routledge.
- Van Parij, P. 2003. "Difference Principles". In *The Cambridge Companion to Rawls*, pp. 200-240.
- Van Zyl, P. 1999. "Dilemmas of Transitional Justice: The Case of the South African Truth and Reconciliation Commission," *Journal of International Affairs*, 52/2, pp. 647-67.
- Velleman, D. 2006. "The Self as a Narrator". In *Self to Self. Selected Essays*. Cambridge University Press, pp. 203-223.
- _____. 2009. *How We Get Alone*. Cambridge University Press.

- Verdeja, E. 2006. "Reparations in Democratic Transitions". *Res Publica*, 12, pp. 115-136.
- _____. 2010. "Official Apologies in the Aftermath of Political Violence". *Metaphilosophy*, 41/4, pp. 563-581.
- Voorhoeve, A. 2005. "Incentives and Principles for Individuals in Rawls's Theory". *Ethics and Economics*, 3/1, pp. 1-7.
- Waldron, J. 1992. "Superseding Historic Injustice". *Ethics*, 103, pp. 4-28.
- Walker, M. 2010. "Truth Telling as Reparations". *Metaphilosophy*, 41/4, pp. 525-545.
- Wall, S. 1998. *Liberalism, Perfectionism, and Restraint*. Cambridge University Press.
- _____. 2003. "Just Savings and the Difference Principle." *Philosophical Studies*, 116/1, pp. 79-102.
- _____. 2007. "Collective Rights and Individual Autonomy". *Ethics*, 117, pp. 234-264.
- Weine, S. 2006. *Testimony After Catastrophe: Narrating the Traumas of Political Violence*. Northwestern University Press.
- Weschler, L. 1989. "Afterword". In *State Crimes: Punishment and Pardon*, Aspen Institute.
- Wenar, L. 1995. Political Liberalism. An Internal Critique. *Ethics*, 106/1, pp. 32-62.
- Wiens, D. 2012. "Prescribing Institutions without Ideal Theory". *The Journal of Political Philosophy*, 20/1, pp. 45-70.
- Williams, A. 1998. "Incentives, Inequalities, and Publicity". *Philosophy and Public Affairs*, 27, pp. 225-247.
- Wolff, J. 1998. "Fairness, Respect, and the Egalitarian Ethos". *Philosophy and Public Affairs*, 27/2, pp. 97-122.
- _____. 2010. "Fairness, Respect, and the Egalitarian Ethos Revisited. *Journal of Ethics*, 14/3, pp. 335-350.
- Wood, A. 2008. *Kantian Ethics*. Cambridge University Press.

Zalaquett, J. 1998. "Moral Reconstruction in the Wake of Human Rights Violations and War Crimes". In Moore, J. *Hard Choices: Moral Dilemmas Relating to Humanitarian Intervention*. Rowman & Littlefield Publishers.

_____. 1999. "Truth, Justice and Reconciliation: Lessons for the International Community". Zalaquett, J. *Comparative Peace Processes in Latin America*. Stanford University Press.