

LOGOS

The Cornell Undergraduate Journal of Philosophy

Is Foucault's Theory of Power Incompatible with Normative Criticism? _____

Michael Law-Smith, University of British Columbia

Who are 'Women'? Defining the Term to Best Serve Feminist Goals _____

Naseeb Bolduc, University of British Columbia

Shortcomings of Fairness for Political Obligation _____

Travis Quigley, University of Virginia

Sexual Consent and Deception: A Reply to Dougherty _____

Ben Atlas, University of Cambridge

Husserl's 'Science of Pure Consciousness' _____

Daire Boyle, Maynooth University

Volume XIII - Spring 2017

Logos : λόγος

Volume XIII • Spring 2017

Copyright © 2017
Logos: The Undergraduate Journal of Philosophy
orgsync.com/72275/chapter
Cornell University Ithaca, NY

An independent student publication

Logos: The Undergraduate Journal of Philosophy, an independent student organization located at Cornell University, produced and is responsible for the content of this publication. This publication was not reviewed or approved by, nor does it necessarily express or reflect the policies or opinions of, Cornell University or its designated representatives.

Mail:
Logos: The Undergraduate Journal of Philosophy
c/o The Sage School of Philosophy
218 Goldwin Smith Hall
Cornell University
Ithaca, NY 14853

Email:
journal.logos@gmail.com

Logos : λόγος

Volume XIII • Spring 2017

Editors-In-Chief:

Noam Weinreich

Club President:

Kwesi Kwapong

Vice President:

Arlinda Shehu

Treasurer:

Rohin Garg

Chair of Advertising:

Victor Chen

General Editorial Staff:

Janilya Baizack

Toby Dresdner

Rebecca Ebiana

Inyoung Hong

David Kahn

Henry Klausner

Haoxuan Ma

Madilynn Madigar

Joshua Nathanson

Dominique Nikolaidis

Tianyang Xu

Mcrid Wang

Faculty Advisor:

Harold Hodes

Contents

Editors' Introduction	9
Is Foucault's Theory of Power Incompatible with Normative Criticism <i>Michael Law-Smith</i> <i>University of British Columbia</i>	11
Who are 'Women'? Defining the Term to Best Serve Feminist Goals <i>Naseeb Bolduc</i> <i>University of British Columbia</i>	27
Shortcomings of Fairness for Political Obligation <i>Travis Quigley</i> <i>University of Virginia</i>	39
Sexual Consent and Deception: A Reply to Dougherty <i>Ben Atlas</i> <i>University of Cambridge</i>	57
Husserl's 'Science of Pure Consciousness' <i>Daire Boyle</i> <i>Maynooth University</i>	67

Editors' Introduction

The staff of *Logos* is proud to present the thirteenth volume of Cornell University's undergraduate journal of philosophy. After carefully considering the submissions we received over the past year we have selected an exemplary set of five articles chosen for their creativity, cogency, and depth of philosophical inquiry.

This year's selection pool was full of quality submissions, and we received inquiries from over fifty undergraduates situated across the English-speaking world. All of the papers contained within this volume were carefully reviewed and selected because of their exceptional quality and varied subjects. The thirteenth volume of *Logos* features papers whose topics fall under the headings of Foucault, metaphysics of gender, political philosophy, ethics of consent and Husserl. We are delighted to be able to publish such a broad set of articles while bringing the best new undergraduate work to public view.

We would like to thank and acknowledge the authors of our chosen submissions: Michael Law-Smith for his submission entitled "Is Foucault's Theory of Power Incompatible with Normative Criticism?," Naseeb Bolduc for his submission entitled "Who are 'Women'? Defining the Term to Best Serve Feminist Goals," Travis Quigley for his submission entitled "Shortcomings of Fairness for Political Obligation," Ben Atlas for his submission entitled "Sexual Consent and Deception: A Reply to Dougherty," and Daire Boyle for her submission entitled "Husserl's 'Science of Pure Consciousness'."

We are grateful to the Student Assembly Finance Commission whose funding supports *Logos*. We are deeply indebted to the staff of the Sage School of Philosophy, particularly our advisor Harold Hodes, as well as Pamela Hanna and Dorothy Vanderbilt, for assisting with publication, the Life Raft Debate, and the day-to-day of running the journal; and to our undergraduate staff without whom none of this would be possible.

Noam Weinreich
Editor-in-Chief

Is Foucault's Theory of
Power Incompatible with
Normative Criticism?

Michael Law-Smith
University of British Columbia

Michel Foucault's books *Discipline and Punish* and *The History of Sexuality [Volume 1]* are praised for offering an original—and discomfiting—critique of contemporary social practices, institutions, and values. In these “genealogical” works, Foucault argues that power and knowledge are inseparably linked. For Foucault, the manner in which discourses—including, in particular, scientific and humanistic discourse—produce knowledge is never power-neutral. Foucault's histories of the discourses surrounding punishment and sexuality—and of these discourses' relation to institutions such as schools and hospitals—are intended to reveal the normalizing and self-shaping effects of “modern power.” Ultimately, Foucault's genealogies are meant to encourage readers to challenge the direction and outcomes of such discourses. In short, Foucault's genealogies are meant to offer a normative basis—i.e., moral and political reasons—for criticizing such discourses and the institutions they promote.

In response to these ideas, three writers—Jurgen Habermas, Nancy Fraser, and Charles Taylor—have argued (in separate works) that Foucault's theory of power undermines his normative project. While each writer focuses on different aspects of Foucault's genealogies, they are alike in arguing that Foucault's theory of power is inconsistent with normative critique. For Habermas, Fraser, and Taylor, Foucault's theory of power renders his normative claims incoherent.

This essay explores a possible response to the above objection. I argue (1) that Foucault's normative claims can be reconciled with his theory of power if a “perspectivist” view of normative justification is adopted and (2) that Foucault can be plausibly interpreted as holding just such a view. The particular perspectivist approach that I explore draws on an interpretation of Nietzsche defended by Bernard Reginster. In this account, normative critiques¹ of the kind advanced by Foucault can be justified, not by recourse to universal criteria, but by recourse to criteria internal to the discourse that they criticize.

I do not argue that Foucault explicitly adopted perspectivism, or even argue that all aspects of his work are consistent with perspectivism. Nor do I argue that perspectivism is the best theory of normative justification. Finally, I make no attempt to defend Foucault's theory of power from other objections. My argument is simply that perspectivism offers a possible solution to the objections described above and, moreover, offers a solution that Foucault could plausibly adopt.

This essay is in three parts. Part 1 describes Foucault's critical histories

¹ By “normative critique”, I mean normative claims in the form of criticism—e.g., “institution Y is bad, so X should not support institution Y” or “belief Z is problematic, so X should not hold belief Z.”

of the discourses surrounding punishment and sexuality, explains why Foucault conceives of these histories as “genealogies”, and outlines how that conception amounts to normative criticism. Part 2 presents Habermas, Fraser, and Taylor’s arguments that these genealogies embody a normative contradiction. Finally, Part 3 argues that a perspectivist reading of Foucault provides a possible route to overcome the normative contradiction.

FOUCAULT’S GENEALOGIES

Foucault describes *Discipline and Punish* and *The History of Sexuality* as “genealogies”. Genealogy, for Foucault, is the historical analysis of “the present”—that is, the history of how our current beliefs or practices came to be.² The distinguishing feature of genealogy, for Foucault, is its rejection of teleological interpretations of the past. Instead of imposing “suprahistorical” meanings or necessary progressions onto past events, Foucault aims to offer a “grey, meticulous, and patiently documentary” account of the forgotten “slips”, “reversals”, and “accidents” that shaped our institutions and beliefs.³ Foucault’s genealogies are thus meant to challenge the common perception that our current values and practices are inevitable.

Of course, a value or practice’s contingency is not itself a reason to reject or even criticize the value or practice. However, it is a necessary premise of any critical claim: if a value or practice is necessary, critique is pointless. The next step in a critique—giving reasons to challenge that which is contingent—is also taken in Foucault’s genealogies. Foucault is clear that his genealogies are meant to “problematize” values and practices.⁴ For Foucault, genealogy aims to show how values and practices became problematic in the past and how they can be problems, or “dangerous”, in the present—specifically due to the “modern power” they yield.⁵

Discipline and Punish illustrates this method. It opens with an analysis of two historical accounts of punishment. The first is an 18th century witness account of a brutal torture and public execution. The second is a 19th century prisoner’s timetable, which describes a strict and meticulous daily schedule. The question that Foucault addresses in the context of these accounts is how the social norms surrounding punishment transformed from supporting public

² Foucault, Michel. “Nietzsche, Genealogy, History” in *The Foucault Reader*, ed. Paul Rabinow (New York: Vintage Books, 2010), 89.

³ Foucault, “Nietzsche, Genealogy, History”. 96.

⁴ Foucault, Michel “What Our Present Is” in *Foucault Live: Collected Interviews, 1961-1984*, ed. Lotringer, Sylvère, (New York: Semiotext(e), 1996), 94.

⁵ Foucault also uses the term “problematization” to refer to how certain practices arise as a response to topics or other practices becoming problematic. The rise of the prison, for example, was a response to problems that were being noted in punishment practices.

torture to calling for private discipline. Foucault writes that the “Enlightenment narrative” of this transformation is that punishment became more humane. However, Foucault believes that this shift—which he characterizes as a shift from punishing “the body” to punishing “the soul”⁶—did not necessarily lessen the criminal’s subjugation. For Foucault, the modern practice of observing and recording every detail of a prisoner’s life exercises a new kind of subjugating power—that is, “disciplinary power.” In this view, the detailed surveillance, examination, and evaluation that occur in modern prisons enforce—and normalize—a rigid, scientific conception of identity and action. As prisoners’ actions and thoughts are observed, recorded, and self-confessed, the category of “the criminal” as a type of person is reified, and subcategories of specific prisoners’ “nature” become a type of knowledge. Under panoptic surveillance—that is, surveillance in which prisoners always feel as though they could be watched—prisoners begin to discipline their own behavior through the norms inherent to this body of knowledge. Further, for Foucault this self-shaping power is not limited to prisons: it is dispersed so as to create “a society of normalization” in which a range of institutions (notably schools and hospitals) employ surveillance-disciplinary techniques.⁷ Foucault warns that if we view these practices and the knowledge that they produce as inevitable we will fail to see the power they exert over how we conceive of ourselves and how we act.

*The History of Sexuality [Volume 1]*⁸ offers a different illustration of the genealogical method. In this book, Foucault questions the “repressive hypothesis”—the idea that the 17th to 19th century was a period in which sexuality was repressed.⁹ Foucault argues that while some sexual acts may have been repressed, the discourse on sexuality flourished during this time. During the repressive period, sexuality became a topic of knowledge—specifically, a fundamental component of the emerging human sciences. The new psychiatric and medicinal study of sex developed sexual categories and sexual “facts” that continue to dominate contemporary thinking about sex.¹⁰ Central to this proliferation of sexual discourse, for Foucault, was the emergence of the notion that there is a “truth” to one’s sexuality—roughly, that every individual has a true sexual nature.¹¹ For Foucault, this notion was a social construction, not a scientific discovery. Just like the surveillance-discipline techniques discussed in *Discipline and Punish*, the belief in a scientific truth about sexuality has influenced—has had a “power” over—how we think about ourselves and act in society.

⁶ Foucault, “What Our Present Is” p. 16

⁷ Foucault, Michele “Two Lectures” in *Power/Knowledge: Selected Interviews and Other Writing, 1972-1977*, Trans: Colin Gordon. (New York: Pantheon, 1980), 44.

⁸ In this essay, “*The History of Sexuality*” always refers to Volume 1.

⁹ Foucault, *The History of Sexuality, Vol I*. Trans. Robert Hurley. (New York: Vintage, 1990), 17.

¹⁰ Foucault, *History of Sexuality [Volume 1]*, 53.

¹¹ *Ibid.*, 69

As these examples show, Foucault's genealogies are meant to explain how we have arrived at the contemporary conception of ourselves and our institutions. And that conception, for Foucault, is intimately connected to the history of social power. In this understanding, power relationships always constitute a field of knowledge, and knowledge necessarily exists within a power relation.¹² In other words, instances of modern power always give rise to truth claims and, furthermore, truth claims always have effects on individuals. This view rejects traditional theories of knowledge, which suppose that knowledge and truth are independent of the social world. For Foucault, knowledge is dependent on the norms and criteria of a given discourse, and so is embedded within social interactions. From this perspective, the creation of "truth" is always an exercise of power: establishing the truth determines which beliefs are legitimate, which, in turn, establishes norms for thinking about the self. Given the fixed relationship between knowledge and power, Foucault concludes that it is impossible to achieve a society free of power.¹³

Foucault's approach also rejects traditional theories of power. Power is traditionally conceptualized as the control that one person or group has, in a top-down direction, over another person or group. For Foucault, "modern" power can also be exercised from the bottom-up. More broadly, power, in this view, is diffused throughout society in everyday relationships—e.g., parents and children, teachers and students, friends—that shape behavior. Individuals self-discipline themselves as they become aware of the norms implicit to the framework of knowledge by which they are being judged in these relationships.¹⁴ Thus understood, power is not always exercised onto the subject from above (e.g. from the state): it can also be exercised from within the subject, through self-discipline. Conceived in this way, power does not always restrict those who are subject to it; power can be productive, i.e., power can empower those subject to it.

Foucault's theory of power is not, standing alone, a normative critique. It is an explanation of how we have come to conceive of ourselves and, more broadly, to think about the world in the way that we do. However, as mentioned earlier, by exposing the contingency of our thinking, the theory invites critique. Further, this is an invitation that Foucault (and many of his followers) have taken up. Notwithstanding his occasional claims to neutrality,¹⁵ Foucault appears critical of many of the practices and discourses that he discusses. At the least, Foucault's genealogies of punishment and sexuality seem committed to the view

¹² Foucault, *Discipline and Punish*, 27.

¹³ Foucault, "Truth and Power" in *Power/Knowledge: Selected Interviews and Other Writing, 1972-1977*, 225.

¹⁴ *Discipline and Punish*, 131.

¹⁵ See, for example, Foucault's interview "Clarifications on the Question of Power" in *Foucault Live: Collected Interviews, 1961-1984*, 262.

that (a) there are aspects of modern prison systems that are bad, and (b) there are aspects of how we conceptualize sexuality that are bad. More generally, Foucault's rhetoric—for example his description of particular discourses as “problematic” or “dangerous”—is not neutral. Finally, even if Foucault maintained a consistently descriptive approach to the topics he discusses, it would be important to ask—as this essay asks—whether the theory underlying Foucault's work allows for normative critique. Foucault is an influential social theorist, and many of his readers are interested in social reform.

FOUCAULT'S NORMATIVE INCOHERENCY

This section examines three arguments that seek to show that it is impossible, or at least incoherent, to adopt Foucault's theory of power and, at the same time, to make normative critiques. The first argument, by Habermas, contends that the relativism of Foucault's theory of power is self-referentially defeating.¹⁶ The second, by Fraser, argues that Foucault's normativity presupposes the liberal normative framework that it is meant to challenge.¹⁷ Finally, the third argument, by Taylor, holds that by defining power in neutral terms, Foucault forfeits any reason to resist power.¹⁸

In Lectures IX and X of *The Philosophical Discourse of Modernity*, Habermas argues that Foucault's “relativism without recourse” undermines the possibility of adopting a critical position. Foucault's theory is relativistic because it denies any universal criteria for truth. This relativity, Habermas argues, leaves Foucault's genealogies “cryptonormative”—by which he means, in part, that the presence of normativity within them is mysterious and ultimately unwarranted.¹⁹ If Foucault's theory is correct—that is, if all knowledge is a consequence of power relationships—then Foucault's own normative criticisms are “nothing more than the effects of power.”²⁰ Without knowledge or criteria that can stand outside of power—criteria that could, for example, compare two power regimes—Foucault cannot evaluate a regime. Habermas thus concludes that Foucault's critique is self-referentially defeating.

Nancy Fraser's objection focuses on a related inconsistency that arises when Foucault attempts to make normative critiques. Fraser argues that Foucault's critiques presuppose the classical liberal normative framework that his

¹⁶ Habermas, Jürgen. *Lectures IX and X in The Philosophical Discourse of Modernity: Twelve Lectures*. (Cambridge, Mass: MIT Press, 1987).

¹⁷ Fraser, Nancy. “Foucault on Modern Power: Empirical Insights and Normative Confusions” *Praxis International* 1.3 (1981): 272-287.

¹⁸ Taylor, Charles. “Foucault on Freedom and Truth.” in *Foucault: A Critical Reader*, ed. David Couzens Hoy. (Oxford: Blackwell, 1996).

¹⁹ Habermas, 282-286.

²⁰ *Ibid.*, 281.

conception of power is meant to transcend. The liberal framework determines whether an instance of power is “legitimate” by asking whether a subject’s rights were transgressed.²¹ However, on Foucault’s account of modern power—as a dispersed, ground-up, and productive phenomenon that can be exercised *through* a subject and not *onto* a subject—power is not a “relationship” of which this question could be asked. If modern power is neither emitted from a central source nor onto a subject, the liberal framework which asks whether X legitimately has power over Y does not apply. Yet Foucault, Fraser argues, appeals to this framework when criticizing institutions of modern power, in particular when he describes them as “dominating” or “subjugating” individuals.²² Indeed, Fraser argues that underlying Foucault’s worry about disciplinary power and sexual discourse is a worry about how they control our autonomy—which again appeals to the liberal framework. Fraser concludes that by failing to provide a competing (non-liberal) normative framework through which to evaluate modern power, Foucault’s normative criticisms have no foundation.

In a similar vein, Charles Taylor argues that Foucault’s notion of power undermines his normative claims. Taylor argues that the concept of power is only coherent if power is understood as in some way restricting liberty or concealing truth.²³ As we have seen, Foucault defines power more broadly; specifically, Foucault supposes that power can be a productive force. Yet, Taylor urges, if power does not hinder freedom or disguise truth, what is to be gained by overcoming it? Taylor concludes that Foucault’s genealogies are paradoxical and nihilistic. Paradoxical because Foucault identifies “problems”, but it is unclear on what ground these things are actually problems. Nihilistic because even if the label “problem” is justified, Foucault’s arguments suggest that escaping a problem would require escaping modern power, but, as explained earlier, Foucault thinks that such escape is impossible.²⁴

RECOURSE TO STRONG PERSPECTIVISM

This section argues that Foucault, viewed through a perspectivist lens, can avoid the normative contradiction identified by the above critics. I begin by explaining the perspectivist position, specifically the “strong” version of perspectivism found in Nietzsche’s work. I then argue that Foucault’s work supports, or at least is consistent with, this position. Finally, I explain how strong perspectivism allows Foucault to avoid the normative contradiction identified by Habermas, Fraser, and Taylor.

Strong perspectivism (hereafter “perspectivism”) is a theory of

²¹ Fraser, 280.

²² *Ibid.*, 182.

²³ Taylor, 90.

²⁴ *Ibid.*, 69.

justification that Bernard Reginster attributes to Nietzsche.²⁵ According to perspectivism, the criteria for justifying beliefs are always intrinsic to particular perspectives (that is, roughly, to particular normative outlooks).²⁶ Perspectivism is therefore anti-essentialist in that it assumes that there are no criteria by which different perspectives can be compared and evaluated. However, perspectivism holds that it is possible to justify beliefs from *within* a perspective. Justification within a perspective is a matter of applying the justification criteria inherent to that perspective. Consider utilitarianism as a perspective. As a moral system, utilitarianism has a criterion for determining what is morally right—for Bentham, this criterion is the “greatest happiness principle”. Within the perspective of Bentham’s utilitarianism, therefore, a normative claim is justified if it maximizes happiness. At the same time, the conclusion that a claim is justified within utilitarianism does not entail or require justifying utilitarianism itself.

Perspectivism thus attempts to position itself between subjectivism and universalism. Whether individuals’ beliefs or actions are justified is a matter of whether those beliefs or actions cohere with the internal justification principles found within the perspectives adopted by those individuals; justification is not merely a matter of a subject’s tastes or preferences. On the other hand, while justification criteria exist within a perspective, there are no universal criteria by which different perspectives can be evaluated. Perspectivism is epistemologically and normatively “local”: it localizes the requirements for a justified belief or action.

For a perspectivist critique to motivate action, it must appeal to reasons which the critique’s recipient recognizes as valid. For example, to motivate people to transform their conception of sexuality, Foucault must appeal to principles inherent to the current discourse surrounding sexuality (such as the idea that we ought generally to be able to express ourselves without restriction). As Reginster notes, this aspect of perspectivism aligns with Bernard Williams’ “internalism” about practical reasons.²⁷ Williams holds that individuals only have reason to act if that action serves their “subjective motivational set.” Through a perspectivist lens, we could interpret Williams as saying that normative claims can only motivate action if they appeal to considerations intrinsic to the perspective from which they are made.

Thus understood, perspectivism allows for the possibility of critiquing a particular perspective without committing oneself to that perspective. For example, X might criticize Y, a vegetarian, for failing to stick to their vegetarian

²⁵ Reginster, Bernard. “Perspectivism, Criticism and Freedom of Spirit.” *European Journal of Philosophy* 8.1 (2000): 40-62. Web.

²⁶ *Ibid.*, 40.

²⁷ Reginster, p. 44

creed—notwithstanding that X is not vegetarian and does not even believe in vegetarianism. Similarly, X might criticize vegetarianism generally by arguing that fidelity to vegetarian principles requires not just that one stop eating meat, but that one not use animal products in any way. In other words, perspectivist critiques are always “internal”.²⁸ A critique is internal if it evaluates a perspective’s content (its beliefs, actions, practices, etc.) by the perspective’s own criteria. An external critique, by contrast, evaluates a perspective’s content by criteria endorsed by another perspective—or by criteria that are applicable to all perspectives—that is, by “universal” or “objective” criteria.

As mentioned, Foucault does not describe himself as a perspectivist. However, he can plausibly be interpreted as a perspectivist. Discussing his genealogical method, Foucault explicitly states that, according to his theory, knowledge is located within a perspective: “the final trait of effective history [i.e. genealogy] is its affirmation of knowledge as a perspective.”²⁹ Foucault expresses a yet clearer perspectivist leaning in the essay “Truth and Power” when he writes that the “effects of truth are produced within discourses which in themselves are neither true nor false.”³⁰ Later in the essay, Foucault elaborates:

“The important thing here, I believe, is that truth isn’t outside power, or lacking in power: contrary to a myth whose history and functions would repay further study, truth isn’t the reward of free spirits, the child of protracted solitude, nor the privilege of those who have succeeded in liberating themselves. Truth is a thing of this world: it is produced only by virtue of multiple forms of constraint. And it induces regular effects of power. Each society has its regime of truth, its ‘general politics’ of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements, the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true.”³¹

Here, Foucault appears to support the salient features of perspectivism. He argues that different “regimes of truth”, that is, different perspectives, contain distinctive “forms of constraint”. These constraints are criteria that separate truths from falsehoods, and as such delineate the accepted “techniques

²⁸ *Ibid.*, 42.

²⁹ Foucault, “Nietzsche, Genealogy, History”, 90.

³⁰ Foucault, “Truth and Power”, 118.

³¹ *Ibid.*, 131.

and procedures” for producing knowledge within that system. Importantly, Foucault does not argue that truth claims are misplaced; his goal is only to show that they are not power-neutral. Foucault’s focus is the “effects” of truth claims, not their validity. Expressed in perspectivist terms, we could say that Foucault is attempting to understand how the process of truth-creation within a perspective affects individuals.

Foucault’s normative claims (or the normative claims of a social theorist who wishes to adopt Foucault’s theory of power), interpreted as perspectivist claims, are not vulnerable to the normative contradiction identified by Habermas, Fraser, and Taylor. Interpreted as perspectivist claims, Foucault’s normative claims can be understood as urging readers to assess practices in light of their own perspective’s principles and Foucault’s empirical findings. On a perspectivist reading, Foucault’s normativity is found in its *rhetorical* power to have readers evaluate whether the practices that they have accepted as valid are supported by the perspective they have accepted. Stated differently, Foucault is not claiming that “belief X in perspective Y is wrong,” but rather that “it is wrong to hold belief X within perspective Y.”

The perspectivist thus aims to show that a belief or practice is irrational for specific agents because this belief is inconsistent with other principles in those agents’ perspective, not because the belief is wrong in itself. This method requires that the perspective in question contains a principle of internal consistency. In theory, therefore, this approach cannot be applied if the relevant perspective permits contradictory principles or practices. However, this limitation does not apply to the present argument as internal consistency is an overarching principle of the Enlightenment perspective on which Foucault and his disciples focus.

Perspectivist critique, as described above, is consistent with the contingency-revealing core of Foucault’s genealogies. For a perspectivist, the aim of normative critique is not to propose a universal normative ideal—an ideal that everyone should adopt—but, more modestly, to assess a society’s beliefs and practices against its own ideals. This approach is compatible with these ideals being socio-historical constructions. The perspectivist avoids the critics’ “normative contradiction” because the perspectivist accepts the relativity of normative critique.

Consider the modern, and nearly unanimously endorsed, ideal of autonomy. Broadly speaking, agents are autonomous if their actions are self-governed, rather than externally governed. For someone who adopts this ideal, Foucault’s genealogies provide clear reasons to criticize contemporary views about punishment and sexuality. We could imagine such a critique

along the following lines:

- 1) In general, autonomy is an ideal (i.e. value) of modern society.
- 2) If autonomy is an ideal, then, from the dominant perspective of modern society, it is a criterion by which we can assess current practices—that is, we can ask whether current practices promote autonomy.
- 3) The ways in which modern human sciences have constructed the image of a person has led to practices of punishment and sexuality that are inconsistent with the ideal of autonomy.
- 4) Therefore, we have a normative reason to criticize modern human practices (and the human sciences that support them).

To support this form of normative criticism, Foucault does not have to justify the ideal of autonomy itself. The third premise generates a normative outcome only when viewed through an autonomy-promoting perspective. Viewed from a different perspective, Foucault's analysis is normatively neutral.

A possible objection to the idea of perspectivist critique is that in cases where a perspective, as conventionally understood, appears to contain contradictory beliefs, the perspective provides no criteria for determining which belief should prevail. Even in cases where the conflict is between a principle and a practice (for example, the principle of autonomy and the practice of punishment), there appears to be no reason that the principle should outweigh the practice. A full response to this objection (which applies to perspectivism generally, not merely the Nietzschean version explored here) is outside this essay's scope. Briefly, however, one possible response is that perspectivist critics must attempt to determine which principles, etc., are "central" or "core" to a perspective, and then to use these principles, etc., to assess other principles, etc. This task, in turn, basically involves determining which principles, etc., are regarded by participants that hold the perspective, in their considered reflection, as central. A principle is at the core of a perspective if it is regarded as such by most participants in that perspective—if it is such that, if it were eliminated, the perspective would cease to be recognizable as the same perspective. Of course, it is possible that the participants may be unable to agree on such principles. However, the principle to which Foucault's critique appeals—autonomy—is widely agreed to lie at the core

of the enlightenment perspective that, Foucault rightly assumes, is shared by most of his readers.³²

Interpreted as a perspectivist, Foucault therefore has a response to Habermas' objection. Habermas' argument that Foucault's genealogies are "relativist" and thus "cryptonormative" is an argument about the justification criteria for normative claims. Habermas argues that Foucault cannot hold a normative position because Foucault's relativism precludes him from distinguishing between justified and unjustified claims. Yet on a perspectivist interpretation, Foucault's normative criticisms are to be judged by the principles internal to the perspective that he is exploring. For example, Foucault's critique of modern incarceration-surveillance techniques can be assessed, from within the same "liberal" perspective, against the view that we ought to increase surveillance for security benefits. From within this liberal perspective, the better position is determined not by appeal to universal reasons, but by appeal to reasons that this liberal perspective already accepts—such as autonomy, equality, privacy, and safety.

Thus understood, Habermas' worry is over Foucault's justification for the liberal criteria themselves. However, providing such a justification is neither Foucault's project, nor a project that Foucault thinks is even possible. Foucault can agree, then, with Habermas' claim that he is unable to appeal to criteria outside of power/knowledge regimes in order to compare different regimes. However, Foucault can still evaluate the practices of a particular regime against its own principles. As a perspectivist, Foucault can—contrary to Habermas' objection—distinguish between justified and unjustified normative claims. Rather than trying to motivate us to "become other than we are" for ahistorical reasons, Foucault attempts to find reasons that we have already accepted—reasons within our present society—to transform ourselves.³³

Fraser's basic objection is that Foucault presupposes the liberal principles that he is attempting to critique. Foucault, viewed as a perspectivist, can respond that in order to critique—which, for the perspectivist, means to internally critique—it is necessary to appeal to at least some of the principles held by the position being critiqued. Foucault does not "presuppose" a liberal framework, but simply accepts that a proper critique of liberal principles—that is, a critique which coheres with the implications of his theory of power—*must* itself be situated within a liberal framework. So it is not an objection to Foucault to say, as Fraser says, that the reason readers find Foucault's description of a society infused with power repulsive is that they value Kantian notions of

³² To be sure, not all of Foucault's readers hold autonomy as a paramount value: for these readers Foucault merely offers a basis upon which they may criticize others for inconsistency.

³³ Foucault, "Nietzsche, Genealogy, History", 88.

autonomy.³⁴ From a perspectivist interpretation, this appeal to Kantian autonomy is not an unjustified presupposition, but simply how genealogy functions as a normative text. Foucault's critique is meant to encourage readers to apply their own normative framework to his historical analysis, and that analysis can generally be expected to be a liberal autonomy-seeking framework.

Nonetheless, Fraser's criticisms highlight an important limit to perspectivist critique. Foucault cannot simultaneously critique principle X while appealing to principle X; for example, he cannot appeal to the ideal of autonomy in order to critique the ideal of autonomy. One response to this limitation is for the perspectivist to appeal to "core" features of the autonomy ideal as a basis for critiquing other features of the ideal. Immanent criticism of this kind must, by definition, be incremental, but, over time, it can lead to a radical transformation of the ideals to which it appeals. It seems likely that Foucault's genealogies are intended, in time, to have this kind of effect. For example, the reader's concept of autonomy may expand to require more features after the reader realizes the effects of the human sciences on his or her previous thinking. More broadly, Foucault could appeal to the notion of autonomy as free self-conception in order to reject the view of autonomy as constituted by the possession of various negative rights. This kind of transformation is compatible with perspectivism. By appealing to certain aspects inherent to the concept of autonomy, Foucault can argue for criticizing or abandoning other aspects of our understanding of autonomy (and of institutions that are predicated on those aspects of autonomy).

Taylor's objections that Foucault is guilty of a "self-referential contradiction" and a "presupposition problem" have already been addressed. Taylor's other objection is that Foucault's position leads to nihilism. For Taylor, Foucault's view that power need not be either liberty-suppressing or truth-disguising removes any reason to resist it. Taylor concludes that Foucault cannot answer the question "why fight?" and, for this reason, is forced into a nihilist stance.

Foucault is indeed committed to a *form* of nihilism. However, this form of nihilism does not preclude Foucault from having justified reasons to act or resist in specific circumstances. Nihilism about normativity, I suggest, can take either of two forms. The first, *diachronic* nihilism, is nihilism over time. The second, *synchronic* nihilism, is nihilism at a particular time. Foucault is committed to the first of these forms (diachronic nihilism), but not the second. And the first form is compatible with normative claims.

Diachronic nihilism rejects the possibility that essential or necessary values, progressions, or beliefs exist over time. A perspectivist interpretation

³⁴ Fraser, "Foucault on Modern Power...", 284.

of Foucault requires that Foucault accept diachronic nihilism. As we have seen, perspectivism holds that perspectives cannot be compared using criteria *external* to those perspectives, and, further, that the content of any perspective is contingent and thus may not persist through time. Interpreted as a perspectivist, Foucault must, therefore, be a nihilist about time-independent truths, beliefs, or values—that is, he must be a diachronic nihilist.

However, Foucault is not committed to synchronic nihilism. A synchronic nihilist holds that not only is it impossible to make normative claims that persist over time, it is also impossible to make normative claims that apply to a particular space-time moment. Taylor appears to charge Foucault with synchronic nihilism; for Taylor, Foucault is disabled from ever saying anything normative. However, Foucault, interpreted as a perspectivist, can hold that while there are no universal, time-independent, reasons to resist modern power, there can be reasons to resist for particular people at particular moments in time. Whether particular individuals at particular moments in time have reasons to resist modern power depends on the principles to which they currently subscribe. To be sure, Foucault's theory rejects the possibility of a power-free society: for Foucault, there is no utopia towards which we should be striving. But this anti-utopianism is consistent with caring about what is happening in our society at the present moment. Foucault's normativity is not centered on comparing the present with a possible future state, but the present with the best version of itself.

To conclude, I have argued that Foucault, viewed through a perspectivist lens, may be able to overcome the normative contradiction identified by Habermas, Fraser, and Taylor. For these writers, Foucault's theory of power is inconsistent with universal justificatory criteria, and so inconsistent with normative critique. However, interpreted as a perspectivist, Foucault's criticisms need not rest on universal criteria. Foucault's normative claims can be justified by appealing to criteria inherent to the perspective he is criticizing. In this way, Foucault's social critique can be reconciled with his unflinching anti-essentialism.

BIBLIOGRAPHY

- Ashenden, Samantha and David Owen, *Foucault Contra Habermas*. London: Sage Publications, 1999.
- Foucault, Michel. *Discipline and Punish: The Birth of the Prison*. Trans. Alan Sheridan. New York: Vintage, 1995.
- Foucault, Michel. *The History of Sexuality Vol. I*. Trans. Robert Hurley. New York: Vintage, 1990.
- Foucault, Michel. *Power/Knowledge: Selected Interviews and Other Writing, 1972-1977*. Trans: Colin Gordon. New York: Pantheon, 1980. Print.
- Foucault, Michel, and Paul Rabinow. *The Foucault Reader*. New York: Vintage, 2010. Print
- Fraser, Nancy. "Foucault on Modern Power: Empirical Insights and Normative Confusions." *Praxis International* 1.3 (Oct. 1981): 272-287.
- Habermas, Jürgen. Lectures IX and X in *The Philosophical Discourse of Modernity: Twelve Lectures*. Cambridge, Mass: MIT Press, 1987.
- Lotringer, Sylvère, ed. *Foucault Live: Collected Interviews, 1961-1984*, 2nd ed. New York: Semiotext(e), 1996.
- Kelly, Michael, ed. *Critique and Power: Recasting the Foucault/Habermas Debate*. Cambridge, Mass: MIT Press, 1994.
- Reginster, Bernard. "Perspectivism, Criticism and Freedom of Spirit." *European Journal of Philosophy* 8.1 (2000): 40-62. Web.
- Taylor, Charles. "Foucault on Freedom and Truth." in *Foucault: A Critical Reader*, ed. David Couzens Hoy. Oxford: Blackwell, 1996.

Who are ‘Women’?
Defining the Term to Best
Serve Feminist Goals

Naseeb Bolduc
University of British Columbia

Feminists often discuss ‘women’s’ rights, hold ‘women’ only events, and work to end the oppression of ‘women’ that occurs precisely *because* they are women. To achieve this goal, we, as feminist philosophers, must definitively answer: who are women? Many philosophers have attempted to provide a definition that best serves the goal of feminism. Among them are Sally Haslanger, Katharine Jenkins and Natalie Stoljar. In this paper, I will contend that among these three, an altered version of Haslanger’s definition best serves feminism. The first part of my paper will be concerned with the three feminist philosophers and their definitions. After presenting each account of ‘women’, I will provide a brief analysis of the shortcomings of each definition in relation to the goal of feminism. In the second part of my paper I will provide my own altered version of Haslanger’s definition, and justify why it best serves feminism. To conclude, I will respond to a potential objection to my new definition.

To clarify: I am not aiming to define the concept of gender as a whole, nor am I attempting to define what *it means* to be a woman (or address the idea that there is no such thing as a woman), as this would be too big a project for my brief paper. I am simply providing an analysis in the search for the most beneficial definition of ‘woman’ among these three philosophers. Also, I will not debate the usefulness or harms of developing a definitive meaning of the word ‘woman’. These are reasonable concerns, but they must be addressed at a different time.

I recognize the many different forms of feminism that vary quite significantly, and the three authors that I will look at may not refer to the same one. I do believe, however, that the unifying purpose of all types of feminism, at all points in time, is to end the oppression of women that occurs specifically *because they are women* (Haslanger, 38). This is the goal I will refer to in this paper, even though Jenkins has a slightly different interpretation of “feminist goal” (which is an issue I discuss later). Importantly, women can be privileged or oppressed based on other factors of their identity such as race, social class or ability. A successful definition of woman will allow room for these intersectional concerns.

The task of defining ‘woman’ is not easy, and there are two common problems that philosophers encounter in their attempts to provide a definition. The first is known as the commonality problem, which is the difficulty (perhaps impossibility) of finding a commonality among all women, from all times, places, and cultures that could count as indicating their gender (Haslanger, 37). The second problem is known as the normativity problem (which is also related to inclusivity). This is the concern that all definitions of ‘woman’ are value-laden, and in any circumstance will marginalize (or exclude) some individuals and privilege others (Haslanger, 37). A suitable definition of ‘woman’ will avoid both problems as much as possible.

I.

I will begin by exploring the Haslanger definition of ‘woman’. Sally Haslanger decides to take on an analytical project. Rather than articulating our current concept of ‘woman’, Haslanger will completely revise the term (34). This revision will be based on what work we want the concept to do for us, theoretically and politically, as feminists. Importantly, (despite briefly discussing gender identity in her paper) Haslanger’s revision of ‘woman’ defines it solely as a social class, based on designated subordinate positions and roles in society (39).

With an analytic project in mind, Haslanger decides to focus on oppression, which she defines as “a structural phenomenon that positions certain groups as disadvantaged and others as advantaged or privileged in relation to them” (39). Haslanger asserts that “oppression is a significant fact around which we should organize our theoretical categories” (46) because according to her, oppression is the single experience that all women, through out all times cultures and places, share and have shared (39). This allows her definition to avoid the commonality problem. Her definition can be paraphrased in the following way:

S functions as a woman in context C iff S is observed or imagined in C to have female reproductive organs; which marks S as someone who ought to occupy certain kinds of social positions that are subordinate; Which contributes to S’s systemic subordination and oppression in C (along some dimension). (42)

There are a few important points I must explain in order to fully understand and appreciate Haslanger’s account. First, she specifies that S *functions as a woman* (rather than *is a woman*) because she acknowledges that gender may not be entirely stable, and that a woman may not always function socially as a woman (42). Secondly, Haslanger specifies that S may be “observed *or imagined*” to have certain bodily features, which is an important factor in the discussion of trans-women. The issue of including trans-women in this definition will be discussed further in my paper.

Haslanger also explains that “bodily features presumed to be evidence of a female’s biological role in reproduction” may be indicators of other characteristics that this hierarchical system uses to justify the treatment of women (43). By this she means that if one is assumed to be a female, one may also be assumed to be irrational, emotional, and bad at math (for example). It is not necessarily the case that oppression is based upon one’s reproductive abilities alone (although this is an important factor in the oppression of females, which I cannot explain further due to the limited scope of this paper), but rather that these reproductive abilities serve as markers for other characteristics.

The most clear and obvious issue with this definition is the normativity problem. Haslanger's definition excludes and marginalizes women who:

- a) Are not oppressed on the basis of being women
- b) Are not observed or imagined to have female reproductive features

In regards to category "a)" Haslanger admits that she has privileged certain facts as "definitive", and that this marginalizes women who are not oppressed on the basis of being women. Haslanger accepts that these women do not qualify as women under her definition, because since they are not oppressed, they are "not the ones who matter" (46).

In regards to category "b)", Haslanger asserts that "... for most of us there is a relatively fixed interpretation of our bodies as sexed either male or female, an interpretation that marks us within the dominant ideology as eligible for only certain positions or opportunities in a system of sexist oppression" (42). In other words, Haslanger contends that oppression on the basis of being a woman is determined by the perception of our sex. I believe this view aligns fairly well with feminist goals, because women are subordinated based on how they are perceived and not what organs they actually do have. Sexism occurs often when the wrongdoer does not know with certainty that the woman has female reproductive organs. As a result, this view allows for the inclusion of individuals who are not female but assumed to be female, and excludes females who are assumed to be males (who would not experience oppression on the basis of their "womanhood", at least in regard to gender as class).

Katharine Jenkins finds this view problematic for trans-women. Trans people are "people who identify as a gender other than the one to which they were assigned at birth" (Jenkins, 396). She argues that Haslanger's definition contributes to the marginalization of trans-women (who identify as women despite being categorized as male at birth). I will now explore this alternative view.

For clarity I have extracted the relevant aspects of Jenkins' paper into the following argument:

- P1. Haslanger's definition of women as a social class relies on observed or imagined female physical features (linked to reproduction)
- P2. Some trans-women are not observed or imagined to have female reproductive features

C1. Haslanger's definition of excludes some trans-women (396)

P3. Some trans people who are not observed or imagined to have female reproductive features *identify* as women

P4. Trans identities must be respected as entirely valid, and failure to do so is harmful and conceptually linked to forms of transphobic oppression and violence (396)

P5. The goal of feminism is to end the oppression of women (394)

C2. Therefore, trans-women must be included in a successful feminist revision of the term 'women'.

In order to accomplish C2, Jenkins concludes that an amelioration must be made to Haslanger's definition to include gender identity as a sufficient condition for womanhood. Jenkins then proposes a two-part definition (what she calls twin target concepts, 414) that would solve the normativity problem (similar to what she calls the inclusivity problem, in that Haslanger excludes some trans-women). In this conception, an individual would only need to meet one target concept in order to qualify as a woman under her definition (416). The two concepts are the following:

1. Gender as Class

S is classed as a woman within a context C iff S is marked in C as a target for subordination on the basis of actual or imagined bodily features presumed to be evidence of a female's role in biological reproduction.

Jenkins specifies that 'gender as class' refers to "the way that gendered subject positions are defined by dominant ideology" (408), and this half of the twin target concept is simply a condensed version of Haslanger's original definition.

2. Gender as Identity

S has a female gender identity iff S's internal 'map' is formed to guide someone classes as a woman through the social or material realities that are, in that context, characteristic of women as a class.

Jenkins establishes that 'gender as identity' refers to "the way that gendered subject positions are taken up by individuals" (408). Jenkins also says that having a female gender identity involves "having internalized norms of femininity... and taking those norms to be relevant to oneself" (411). These norms must exist within a system of gender as class, and Jenkins notes that the oppressive class position of women plays a role in shaping female gender identity (412).

Gender as identity introduces a new form of oppression experienced by women: internalized or self-oppression. Self-oppression describes the way norms and gender roles in society harmfully influence the way women view themselves and how they choose to live their lives. Jenkins uses an example to illustrate the difference between gender as class and gender as identity: If a woman (we can call Tina) is not given a job, or offered a lower salary because she is a woman, she is being oppressed because she is perceived as a woman, which is an example of class-based oppression. If a woman (we can call Amy) chooses not to apply for a leadership role or a position that is stereotypically given to men because of implicit messages that have been conveyed to her about her value as a woman, she is experiencing self-oppression through her gender identity (414).

Both these forms of oppression are harmful, but I will briefly contend that perhaps they should not be considered equal because the oppression caused by gender identity does, as Jenkins admits, “have some contact” with the sex-based social hierarchy (412). I will continue to use the example that Jenkins provided for clarity. In this example, the oppression that Amy experiences is harmful. However, I contend that Amy only felt like she could not try for the job position that she wanted because Tina’s experience still occurs regularly in society. A person like Amy would only feel a sense of inferiority because she has experienced or known of situations where women are oppressed and subordinated on the basis of being women. This means that in regards to womanhood, as long as there is class-based oppression we will not be able to get rid of identity-based oppression.

Based on this link, I question whether this twin target concept is really accurate or the best option for feminist purposes. Jenkins justifies this twin conception of woman by arguing that assigning gender identity “a secondary or peripheral status” (403) would mean that trans-women are only counted as women in a secondary or peripheral sense, which would marginalize them. While I agree that feminists should not be complicit in the marginalization of trans-women, I also see that the twin target concept may hinder the feminist goal. One way is that the twin target concept may be too inclusive: too many people would fall under the category of ‘woman’, which would ultimately undermine and devalue the meaning of ‘woman’. For example, a person who presents as a male, is perceived as a male, uses male pronouns and overall benefits from male privilege could still fall under Jenkins’ definition of ‘woman’ if they identified as a woman. I do not believe this is an accurate or fair example of a woman: This person would not have to fear domestic or sexual violence, this person would not be denied equal pay or equal opportunities, and this person would not experience other forms of misogynistic injustice. While they may experience self-oppression, this self-oppression would likely never

be reinforced by personal social experiences. For these reasons, this person should not be equated to the women who face sexist adversity and violence at every level of their lives. I contend that the inclusion of this person in the revisionary category ‘woman’ would disparage the (more detrimental) struggles that individuals who are perceived as women face.

Based on this issue of over-inclusion, feminists need a definition that does not rely on gender identity to be a sufficient condition for womanhood, while still avoiding the exclusion of trans-women. One solution could be an approach that uses a completely different strategy. Thus far Haslanger and Jenkins have taken up an analytical project in attempt to revise what we mean by ‘woman’ to best serve feminist goals. An alternative approach, used by Natalie Stoljar, is to define ‘woman’ as a cluster concept.

I should mention that before arriving at woman as a cluster concept, Stoljar discusses essentialism, the concept of a universal and the ways in which ‘woman’ constitutes a class. These discussions are important for the concept of ‘woman’, but are not relevant to my goal in this paper which is to find a definition that most benefits the purposes of feminism. What is important to know is that Stoljar contends that there is no single property that all women share (288), and as a result turns to the idea of a cluster concept. She describes a cluster concept in the following way: “there is a cluster of different features in our concept of woman and in order for an individual to satisfy the concept, it is sufficient to satisfy *enough of*, rather than all and only, the features in the cluster” (282). This approach would allow for diverse women who may typically not be considered ‘women’ to be fit into the definition. Stoljar offers four different features that may serve to indicate membership as a ‘woman’:

- i) Female Sex, and any bodily characteristics associated with the female sex (such as gait or voice quality)
- ii) A feeling of being a ‘woman’ stemming from lived experiences as a woman (such as menstrual cramps, child-birth, sexual experiences as a female)
- iii) Social roles associated with ‘women’ (such as being oppressed, being responsible for child-rearing, wearing typical female dress)
- iv) Self-attributions and the attributions of others, which basically means calling oneself a woman and being called a woman (Stoljar, 284)

Stoljar suggests that an “exemplar” must satisfy at least three of these conditions in order to be considered a woman (284). This approach helps to

solve a number of issues. First, ‘woman’ as a cluster concept allows for the consideration of intersectionality, as factors such as race, social class and ability do not limit an individual’s ability to fit into the definition of ‘woman’. This would also allow us to acknowledge oppression experienced along different dimensions. In addition, this cluster concept helps us circumvent the commonality problem, which is the difficulty in finding a single common feature shared by all women.

The idea of a cluster concept also opens up a variety of issues. First, as Stoljar notes, the phenomenology, role, and attribution of ‘womanness’ are all closely related to having a female sex in the first place (284). Many of the “lived experiences” would also require a female sex, which may lead to exclusion of some individuals who should be considered women. Many trans-women, some intersex individuals (who are oppressed as women) and women with variations in their anatomy may have a difficult time satisfying these conditions.

This approach is also susceptible to the normativity problem. Some women who easily meet all four conditions may be considered better examples of women (285). Similarly, since these four conditions are “revisable” (286), it is not clear how we can be sure that we are using the correct conditions or the correct number of conditions. Also, Stoljar does not justify her reasoning behind requiring that three conditions be met (284), rather than one, two, or even four. These decisions are all value-laden (and at this point seem arbitrary), which ultimately leave the cluster-concept model open to the normativity problem.

Given these issues with the cluster-concept, I believe that an analytic project holds the potential to best serve feminist goals. Since Haslanger’s version excluded many trans-women, and Jenkins’ approach problematically equivocates gender identity-based-oppression with gender as class-based-oppression, I will now attempt to conceive my own definition based on Sally Haslanger’s work.

II.

I have made the following claim in part I of my paper:

1. Equivocating the oppression caused by ‘gender identity’ with ‘gender as class (by finding them to be twin target concepts) is harmful for feminist politics because it undervalues the subordination that women experience in society *because they are women**.

I would like to elaborate, that this would also contribute to the epistemic injustice that women experience as women: A woman's testimony, even regarding her own experiences, is often not given the same value as the testimony of men. To have individuals benefit from male privilege and claim to understand the harms caused by the oppression of women would further add to the depreciation of women's testimony.

I would now like to make a second claim:

2. Denying trans-women's identities harms and marginalizes them but not on the basis of being women.

I make this assertion because this is a harm not exclusive to trans-women: trans-men and non-binary individuals can also experience the denial of the legitimacy of their identities. I do not believe that feminists want to be complicit in the marginalization or oppression of anybody. However, the original goal in defining 'woman' was to eliminate oppression that occurs on the basis of being a woman. I understand that Jenkins defines feminism to be "... a movement to end the oppression of 'women'" (394), but I believe she interprets this to mean every form of oppression that every woman faces. I disagree: the goal cannot be to eliminate *all* the oppression that *all* women face (which would be an overwhelming goal), if it comes at the cost of an effective revision of the term 'woman'. It is clear that the core concern of feminism is the oppression of women that occurs because they are women.

While these claims show that Jenkins' twin target concept may not best serve feminist goals, I do think she was on the right track in her strategy to ameliorate Haslanger's original definition. There is another way in which we can alter Haslanger's definition of 'woman' to include most (almost all) trans-women in a way that is justifiable, and in a way that best serves feminist goals. My solution is the following definition:

S functions as a woman iff S is perceived as being (or wanting to be recognized as) a female; which marks S as someone who ought to (or deserves to) occupy certain kinds of social positions that are subordinate; Which contributes to S's systemic subordination and oppression in C (along some dimension).

Just as with the Haslanger definition, I must explain certain points here for a full understanding. I replaced the first requirement in Haslanger's version:

"S is observed or imagined in C to have female reproductive organs" with *S is perceived as being (or wanting to be recognized as) a female*.

This is a solution to the original problem that Haslanger faced, because

any trans-women who were not imagined to have female reproductive organs were excluded in her definition. My version is broader, allowing the inclusion of a wide range of trans-women. In order to specify the degree to which my definition helps, I will refer to the four scenarios described by Jenkins in which a trans-woman may find herself:

1. S does not present as a woman, S is perceived as a man
2. S presents as a woman, but is perceived as a man pretending to be a woman
3. S presents as a woman, S is perceived as a woman
4. S presents as a woman, S is perceived as a woman but not perceived to have reproductive abilities

Of these four scenarios, my definition allows for the inclusion of any individual in 2, 3, and 4. I would like to specify that “wanting to be recognized as female” (as stated in my definition) does not necessitate an effort to surgically alter one’s body or appearance. I simply mean that there is a message conveyed to society (in any way) that S wants to be considered a woman. The importance of this clarification will soon be clear. First, I must briefly address the individual in scenario 1, who is the sole trans-woman who is excluded from my definition.

This person benefits from the privileges that I described in part I of my paper, and I cannot amend my definition further (thus weakening the strength of the definition for feminist purposes) to accommodate this particular scenario. I accept that my definition fails to include all trans-women, but despite this I still believe my version of ‘woman’ best serves feminist goals.

I modified Haslanger’s definition in another way: I specified that S is marked as someone *who ought to (or deserves to) occupy certain kinds of social positions that are subordinate*. I included the idea of “deserving” a social position specifically with trans-women in scenario 2 in mind. I believe that the subordination that trans-women experience when they are seen as “men pretending to be women” is a consequence of misogyny. This is also why I included all individuals who are *perceived as wanting to be recognized as female*. If S is perceived as wanting to be female, S is seen as accepting and embracing the “inferior” gender role of ‘woman’, which makes trans women like S targets of violence and abuse. The apparent sex of S (as male) does not protect S from the hatred aimed at women.

One may object to my view by returning to the problem of gender identity. I imagine an argument could be made by Jenkins that would look

something like this: Since women face (internal/self) oppression based on the fact that they identify as women, these oppressions are relevant to feminism and the ultimate feminist goal of eliminating oppression of women that occurs because they are women. Since this internalized oppression is relevant to our feminist goal, one should include every person who experiences this oppression in the revised definition of 'woman'.

I agree that self-oppression (on the basis of identifying as a woman) is relevant to feminism and the feminist goal. I also agree that the purpose of feminism includes eliminating self-oppression. I recognize that it is harmful, and I recognize that people who are perceived as men but who identify as women may experience this harm.

However, I did not include gender as identity in my definition because I do not find it to be a sufficient or necessary condition for experiencing systemic class-based oppression. In contrast, systemic class-based oppression is a necessary and sufficient condition for the existence of self-oppression through gender as identity. One form of oppression is more foundational, more harmful and must be addressed before we can ever hope to eliminate the other. Again, including both as equally sufficient conditions would undervalue the class-based subordination that women experience in society *because they are women*. This is not to say that trans-women are not women in the ordinary sense. I am not trying to reflect the way 'woman' should be defined in general, but rather to find a way to define 'woman' that best serves feminist purposes. Gender as class is the root of self-oppression, and allowing ourselves to solve the class problem will ultimately lead to a solution to the self-oppression associated with gender identity.

I understand that this may be interpreted as assigning trans-women a secondary or marginal position as women. I refute this claim, as my new definition based on Haslanger's work does include a majority of trans-women, intersex individuals, and anyone else who is perceived as being (or wanting to be recognized as) a woman. The only potential person given a secondary position is a person who receives social privileges as a perceived man, and this is a conclusion I am willing to accept given my purpose of serving the goal of feminism.

In conclusion, it is apparent that defining 'woman' is not an easy task, and many problems exist within each potential definition, including my own. I hope that my analysis and new altered version of Haslanger's work can contribute in some small way to reaching an incontestable definition for the sake of one day achieving the goal of feminism.

BIBLIOGRAPHY

Haslanger, Sally. (2000) "Gender and Race: (What) Are They? (What) Do We Want Then to Be?" *Nous* 34.1 : 31-55

Jenkins, Katharine. (2016) "Amplification and Inclusion: Gender Identity and the Concept of Woman" *Ethics* 126.2 : 394-421

Stoljar, Natalie. (1995) "Essence, Identity and the Concept of Woman" *Philosophical Topics* 23.2: 261-293

Shortcomings of Fairness for Political Obligation

Travis Quigley
University of Virginia

Fairness (or fair play) accounts of political obligation propose a powerfully intuitive reason to believe we should obey the laws of our state: we treat our fellow citizens, who presumably do obey the law, badly by our disobedience. The canonical formulation of the principle of fair play is given by H.L.A. Hart:

When a number of persons conduct any joint enterprise according to rules and thus restrict their liberty, those who have submitted to these restrictions when required have a right to a similar submission from those who have benefited by their submission. The rules may provide that officials should have authority to enforce obedience...but the moral obligation to obey the rules in such circumstances is due to the cooperating members of the society, and they have the correlative moral right to obedience.¹

Fairness theory has, then, been under debate for over five decades. I, of course, will not fully survey the myriad criticisms and reformulations of the principle over that time. I will instead use the extensive literature as a guide and point of departure for my own criticisms of fairness theory, engaging where possible with contemporary debate, and using prominent critics and defenders of fairness (most notably John Simmons and George Klosko, respectively) to delineate the battle lines, so to speak.

My critique of fairness will come in two parts, which relate but are largely freestanding; these are different paths to rejecting fairness theory. In short, the first, longer branch of criticism evaluates fairness accounts on the traditional terrain: I argue that fairness does not yield sufficient duties to meet the goals of political obligation in modern, liberal states. This argument engages directly with the contemporary literature, and will break up into multiple strains of fairness theory, each of which I argue are more successful than some critics suggest, while ultimately still falling short of their goals. Second, I suggest that fairness theory as usually described appears to leave little room to justify obedience to what I will call quasi-just states - institutions that are significantly flawed, but nonetheless either themselves provide significant benefits, or act as barriers stemming a tide of significant harms. I argue that a successful account of political obligation ought to motivate obedience to quasi-just states in at least some circumstances, and that the lack of capacity of fairness to do so suggests that its justificatory power is insufficiently broad.

The structure of these objections, then, is essentially “even if.” I argue that fairness has mechanical shortcomings displayed in the most comprehensive

¹ “Are There Any Natural Rights?” *Philosophical Review* 64 (April 1955): 185

available applications of the principle, which extend to questions regarding the strength of the principle itself. But, even if these fail, and fairness succeeded in much the way its defenders present it, I argue that it will not justify obedience to quasi-just states that political obligation ought to encompass.

MECHANICAL OBJECTIONS TO FAIRNESS

By “mechanical objections,” I just mean criticisms that focus on whether fairness actually functions in practice - whether it shows political obligation to all the people it claims, in the way it claims (as opposed to a claim against the basic strength of the principle, for instance). The most prominent recent critic of fairness on these grounds, of course, has been John Simmons. The basic structure of Simmons’ argument begins by distinguishing exactly how a fairness claim generates obligation. He defends fairness from the charge, still occasionally raised, that it collapses into a consent account, giving examples of people accepting goods from a collective while never consenting (perhaps explicitly disconsenting) to be bound by it. He thinks that these people are in fact bound by the principle of fair play, and that this mechanism, distinct from consent, accounts for a great deal of our actual obligations to groups.² This relies entirely, however, on active acceptance of the goods in question - while the fairness mechanism is distinct from consent, it is a parallel voluntary action. If you accept this far, Simmons thinks he can easily undermine the utility of fairness for general political obligation: in the easy cases for fairness, people actively choose to accept some good. The key goods in a state (security, etc) are rarely of the appropriate nature, since they are public, “open” goods, that naturally accrue to everyone within the state no matter what they do. In order to meaningfully accept public goods, Simmons introduces choice conditions: that the acceptance must be “knowing and willing” - (terms that are only briefly filled out by Simmons, to which we will return later). In brief, actual citizens are not taken, in most cases, to have knowingly and willingly accepted goods from the state, and any account, Simmons claims, that moves *away* from the voluntarist conception (on which knowing and willing acceptance is integral) becomes wildly implausible.³

This line of criticism will set the framework of this section. I will first consider a recent argument from Edward Song that takes the voluntarism condition, intended as something of a theoretical poison pill, and claims that it actually *is* met in modern states. Then I will move on to the more traditional

² Simmons, A. John, *Justification and Legitimacy: Essays on Rights and Obligations*. Cambridge: Cambridge University Press, 2001. 16-18

³ *Ibid*, 22

traditional claim that the voluntarism requirement is misguided, and examine George Klosko's theory of presumptive benefits.

VOLUNTARISTS

Song, in effect, takes the limited way in which Simmons accepts fairness and runs with it. This involves an empirical claim that citizens of modern states actually do accept public goods willingly and knowingly, but also necessitates clearing up a secondary objection raised by Simmons regarding the application of fairness reasoning to the case of large political collectives: that people not knowingly accepting the benefits of the state as a cooperative venture is perfectly reasonable, because the state is far too large to have the "consciousness of cooperation" endemic to cooperative ventures.⁴ This condition was only briefly stated, and subsequently came under attack. Simmons further articulated the point in a later paper, which specified some examples to ground the idea that a common purpose and spirit of cooperation are necessary for a collective to exist, and thus for fairness claims to apply: we can consider a case where people are bound in a competitive, rule-based group, which Simmons takes as decidedly non-cooperative (the members, of course, are explicitly competing, not cooperating); or a case where people *accidentally* generate beneficial goods through the joint effect of their independent actions, with no intent of doing so.⁵

This poses the question of voluntary participation as a two stage test: first, is the system in question one of genuine cooperation? And second, does a given person accept its goods willingly and knowingly? Song's strategy is to undermine the distinction between the two questions. If a group of people willingly and knowingly decide to participate in a joint venture, then Song claims that the venture simply by that fact is sufficiently cooperative to ground obligation.⁶ The definitions of willing and knowing are important here: willing is simply a brute psychological choice state (as Song points out, it is not strictly necessary for participation to make a person better off in order for that person to willingly participate), while knowing entails "accept[ing] some cooperative good when she knows that she is receiving it, knows that the good is generated by a system of cooperation, and knows that the system entails certain requirements."⁷ If these conditions are met, Song claims, the only necessary subjective conditions for genuine cooperation are already in place, meaning the first condition above is automatically satisfied by an affirmative answer to the second.

⁴ *Ibid.*, 25

⁵ *Ibid.*, 38-39

⁶ Song, Edward. "Acceptance, Fairness, And Political Obligation." *Legal Theory* 18, no. 02 (04, 2012). 219

⁷ *Ibid.*, 215

Given that Song thinks such an affirmative answer can be empirically given, this assertion is quite important, and he generates an example to undergird his assertion that just those two conditions are sufficient to answer any qualms about the cooperative nature of an activity, even in an archetypal example of a competitive, rule governed venture:

Consider, for example, athletic competitions. Such examples are precisely the kinds of activities that Simmons wants to rule out...What matters, however, is not whether athletic competitions are examples of the deepest and most attractive kind of cooperation but whether they are cooperative enough for considerations of fairness to come into play....A tennis player might be praised for her fairness when she overrules the umpire's judgment that her opponent's ball landed outside the lines. Or a soccer player's actions might be called unfair because she cheated and broke the rules. These judgments do not seem to be just a way of talking...Such particular judgments, as well as the general notion of fairness, seem to turn on the same idea: it is wrong for a person to give herself preferential treatment because this is to make an exception of oneself, which shows disrespect to others.⁸

There are a number of problems with this example, limiting its intuitive appeal: first, an athlete's relationship to the rules of a sport seem quite different than a citizen's relationship to the law. While Song is correct to note that athletic competitions are not strictly speaking examples of tacit consent, because the "contract" to which athletes agree is implicit, not just their consenting action, using that interaction as a model of political obligation is questionable. I think the disjunction between the two scenarios is roughly the following: a sport is defined only by its rules. Not following the rules, in an important sense, means you *are not playing the sport*. This is plainly not true of citizenship - breaking even relatively major laws like robbery, murder, or even treason does not mean one is not a citizen, because being a citizen is about being a member of a community, as much about personal history as obedience to law or custom. A tennis player is only playing tennis insofar as she plays by its rules; if she chooses not to, she really is playing some distinct variant. A citizen who breaks the law is still fully a citizen; they are not participating in some slightly different country's cooperation.

Second, Song's point about the natural language of unfairness in sports

⁸ Ibid, 217-218

cheating is, I think, equivocating between how people talk about the *results* of cheating and the *act* of cheating. Consider two cases: in one, a soccer player scores a goal using their hand and subsequently wins a championship, personal prestige, and lucrative endorsement deals. In the second, a football coach attempts to trip an opposing player running down the near sideline, but trips and falls in the process, not affecting the game and becoming an object of mockery. These cases are similar in the structure and intent of the act; in each the person sought advantage for their team against the rules of the game. But I think that, in natural language, we would only intuitively refer to the first as unfair. The football coach clearly did something *wrong* - he broke the rules of the game, going against the spirit and value of the activity. But, while he did attempt to gain preferential results under the rules, the relevant sense of unfairness for cheating in sports seems to be about the *result* of the activity. When we say the soccer player's cheating was unfair, what (I think) we mean is that they did not *deserve* the results that occurred, and that malapportionment of benefits is wrong (unfair). The language congruence is simply due to an overlap between fairness as a term of moral *desert* and fairness as a matter of treating yourself preferentially.

Lastly, consider further my claim about the football coach. It seems to me that the major way in which the attempt to trip the opposing player is wrong is that it goes against the spirit of the game - if the coach had been successful, and that had led to a victory for the team, the winning players would likely feel that it was a hollow victory, lacking the desired meaning (i.e. the feeling of genuine victory under agreed upon rules). The fact that this notion of a collective spirit of the game exists speaks to the largest problem with Song's analysis of athletic competitions: they are, in fact, considerably "thicker" instances of cooperation than political states are, so even if fairness accounts could come to bear within them (as I have questioned), it would not immediately imply that the same could hold true for states. Why is the spirit of cooperation relatively thick in athletic competitions? Because all the participants *do*, in fact, have a common purpose: they all want to win the competition under certain shared conditions; the victory would be meaningless if it were only due to some outrageous technological advantage, for instance. Only one competitor (individual or team) can reach that goal, but common interest of *all* competitors dictates that such a victory be available, which is only possible if everyone at least mostly follows the rules (some rules may be looser than others, but many - like coaches not interfering with players - are essentially inviolable). The other example that Song briefly mentions is competing businesses collectively participating in trade organizations.⁹ This seems vulnerable to a similar analysis: while each business would like to defeat its competitors, all share the ultimate

⁹ Ibid, Fn 28, on 218

goal of making money, and their collective action is toward that end, so defection goes against a relatively thick sense of cooperation. This analysis undermines Simmons' reference to rule based competitions as examples of non-cooperation, but upholds his conclusion and counterexamples - in none of these cases are the benefits gained, for instance, at all accidental. Song mentions that there are many other possible examples that he could muster; I suspect that all could be understood in much this fashion.

The fact that these collectives *do* have some common purpose, which in many cases is readily apparent, is due to another disanalogy between Song's examples and actual states. In both sports and business contexts, the actors accepting collective goods are also the ones *forming* the collective. The athletic competition is nothing but a coordinated group of athletes; similarly for the trade association. In this setting, it makes perfect sense that the willing and knowing participation of the actors would always run alongside genuine cooperation - the cooperative in question would not *exist* otherwise. In that sense, the benefits provided by cooperation are at most readily accessible - if they were not freely desired, the cooperative would never form, or a given company or athlete would not participate. Neither of these aspects are whatsoever similar to political states, which from the perspective of citizens provide their benefits without any pause for participation, and clearly *preexist* each modern citizen, offering only cooperation or disobedience as options. This is where Simmons' use of a first stage test of genuine cooperativeness shows its importance: simply imagine a despotic state, that either will rule inefficiently but benignly for cooperators, or harshly punish the disobedient. On Song's account, the fact that many people would willingly (psychologically) and knowingly (with the knowledge that acceptance of the benefits of the despotism entails that it will expect obedience) participate in such a state automatically makes it a subject for voluntaristic fairness. This is an avoidable result, but only if we ask, independently of the willing and knowing conditions, if the state itself is the *right kind of thing* to qualify as a cooperative. A despotism clearly is not; Simmons thinks that even a fairly just, liberal state is not, but Song has largely avoided the question altogether.

Insofar as we could piece together a response from Song in defense of modern states, it would hinge on using the same data with which he supports the willing and knowing conditions to support the status of America (in particular) as a genuine cooperative. Song cites a range of data suggesting that, overall, Americans are quite supportive of and happy with their government. Simmons might say that this stems from habit and ideology, but Song replies, convincingly enough, who cares? If what we concern ourselves with is subjective mental states, then why are the beliefs we find in the actual world disqualified by their

provenance?¹⁰ The problem with Song's conclusion is twofold: first, the simple fact that people *like* their government does not necessarily mean that they think of themselves as cooperating in it. I like the United Nations in some of the same ways as I like the United States, but I surely do not feel that I am bound to the UN by fairness considerations. Second, to the extent that it *could* be shown that America is a genuine cooperative, it would be hard to conclude much about the strength of our fairness obligations from Song's data about general national approval. If the sense in which I willingly and knowingly accept national political benefits is that I vaguely support the nation, it seems difficult to draw out anything but the weakest possible obligation from that basis. Song acknowledges this concern in his conclusion, but does not give it sufficient weight; if, as I suspect, the weakness of his knowing condition necessitates an extraordinarily weak obligation, it is hard to see how the account could stand up to intuitions about political obligation, even if it cleared all the previous objections given here.

NON-VOLUNTARISTS

The strategy used by Song in the previous section, of course, is not the standard one in fairness defenses of political obligation. Rather than attempting to use fairness within Simmons' framework, theorists more commonly go outside it, arguing that some set of benefits, properly defined, can ground obligation without reference to voluntarism. One reason Song uses the voluntarist strategy is an argument that non-voluntarists will have to distinguish between the receipt of excludable private goods and the receipt of nonexcludable public goods, and (he claims) there is no obvious way for them to do so. The thinking is that the distinction between the two kinds of goods is obvious, and the difference lies in the ease of identifying voluntary acceptance of the former but not the latter type. But, the argument goes, if non-voluntarists dismiss the importance of voluntary acceptance, that ought to imply that there is *no* meaningful difference between the types, raising an intuitive problem.¹¹ There seems to me to be a straightforward reply: non-voluntarists are under no obligation to dismiss that voluntary acceptance is *one* important way to incur political obligation, no more than they would have to argue against binding consent. Non-voluntarists merely accept the argument that states cannot be fully justified as voluntarily accepted institutions, meaning that, to justify political obligation, additional principles are necessary. The important distinction between excludable and nonexcludable goods, for non-voluntarists, is which kind they still have to fight over. If their arguments for nonexcludable goods *also* ended up applying to

¹⁰ Ibid, 220-222

¹¹ Ibid, 212

some set of excludable goods, some parts of obligation being multiply justified is not troubling.

The main contender for a non-voluntarist theory of fairness obligation that I will consider is George Klosko's theory of presumptive goods. In short, Klosko acknowledges that large swathes of government functions - in fact, many of the government's most *important* functions, have not been consented to or voluntarily accepted. But he denies that this undermines state legitimacy, because fairness can bind even without voluntarism in the context of public goods of sufficient significance. Presumptive goods are meant to be those public goods that are "indispensable for acceptable lives." This class is not fully specified, but is taken to at the very least include security, as well as public health, environmental protection, and the like.¹²

This kind of account immediately has some appeal: first of all, it at least at first blush appears to have a better chance of powering political obligations of intuitive strength than voluntarist accounts, since it appeals to external, principled duties, rather than voluntary assent of dubious strength. It also appeals to the paradigm cases of significant cooperation - banding together to protect against invasion, or disease, or natural disaster - that seems like they *should* lie at the heart of political obligation. Simmons, perhaps unsurprisingly, does not find this move convincing. In response to presumptive benefit accounts, he raises an example, which seeks to set his objections to non-voluntarism in the context of a presumptive good, and show that they still apply similarly as in any other case.

The example goes as follows: imagine a small town or neighborhood in the throes of a severe drought. One person chooses to dig a deep well in his backyard, and is hard at work on his individual project. Simultaneously, his neighbors band together to dig a trench from a nearby water supply, providing him with fresh water outside his front door - a notable benefit, Simmons notes, even if he regularly declines to use the water from the trench.¹³ But surely fairness does not bind the independent-minded homeowner to have contributed to the collective project when he had no interest in its completion, and never made any voluntary contribution to the cooperative project. This is true despite the fact that the individualist could not have dug the trench by himself, and that water is indisputably a presumptive benefit.

This example, seemingly presented as a knockdown argument, seems to have some significant problems. First, and vitally, it presumes the independent capacity of the individualist homeowner to acquire water, which

¹² *The Principle of Fairness and Political Obligation*, 40-41

¹³ *Justification and Legitimacy*, 34

cuts against the central assumption delineating presumptive goods - in the case of security, if we simply assume that everyone can separately choose to secure themselves without cooperation, and that accepting protection from the state is simply one option among several, the force of security as a presumptive benefit is entirely removed. The water example is only a genuine representation of Klosko's argument if the *only* way to acquire sufficient fresh water is through cooperation. Simmons may disagree with the empirical premise underlying Klosko's theory, and argue that presumptive goods suggested actually *are* independently available, but here he assumes that condition for the sake of example. Second, the availability of water does not even fit neatly into the public good nature of Klosko's account - though fresh water simply existing outside your door is a good thing, you only make use of the significant benefits of water (drinking, bathing, etc) by voluntary action - it seems much more like the neighbors in the example provided a readily accessible good than a properly public, presumptively beneficial one.

Let's consider a modified, stronger example. Rather than hewing exclusively to security, consider an (fanciful) environmental case: say that there is some rupture in the ground that, if not annually plugged, will spew a cancer-causing smog. Presume that the only way to plug the rupture is through manual labor, and it requires a minimum of eighty people working hard to complete the task, out of one hundred people living in the community. The labor, of course, could be spread across all one hundred people, making each person's task easier, but each year five or ten people beg off with no good reason; they may claim to be injured, or claim that the benefit is not worth it to them, but these are in most cases false - they simply choose not to work. This seems a truer representation of presumptive benefits: the benefit unavoidably accrues to everyone in the community, cannot be achieved independently, and is undoubtedly desired by everyone or almost everyone, but nonetheless there is incentive to freeride on the labor of others. It seems to me that, in satisfying Klosko's requirements, we also satisfy Simmons', who writes that "free riders take the benefits others provide without reciprocating, while at the same time preferring receipt of the benefits at the prescribed price to doing without the benefits altogether or to trying to supply them independently in an alternative fashion."¹⁴ This leaves the door open for someone with truly strange preferences to non-culpably refuse to work on the project, but given how rare such a preference set would be, this aspect of non-universality seems untroublesome (on a larger scale, this might explain why large groups of people in the real world could not be directly compelled by fairness to fight against climate change, given that many people, influenced by their political environment, non-culpably believe that carbon emission-reduction actions go against their preferences).

¹⁴ Ibid, 34-35

Insofar as this appears to be Simmons' main argument against presumptive benefits, a better example seems to me to satisfy it. His other stated concerns, for instance that Klosko's argument is actually a disguised natural duty account of beneficence or charity,¹⁵ might be reformulated to still stand, but as-is appear defused by showing that the presumptive benefit account really *does* rest on receipt of a class of benefits, which must be understood in a very particular way. Once that understanding is achieved, it appears to me that the presumptive benefit argument really does show that refusing to cooperate is *de facto* freeriding, and thus fairness claims can succeed on the original conception of the principle, at least in tightly prescribed circumstances.

If non-voluntarist accounts of fairness are successful thus far, why exactly do they fail? One thing is note is that the successful example I used as a rejoinder to Simmons' claim that presumptive benefits cannot generate fairness duties is a very particular case. Its main conceptual characteristic is that the freeriders refuse to contribute toward a good which they presumptively (and actually) want, which is vital enough to the wellbeing of everyone in the community that the remaining members are essentially forced to take up the slack and work harder. Fairness theorists often couch the wrongdoing of freeriders in terms of disrespecting the other members of the community, which is fair enough. But we should be careful about what exactly the substance of that disrespect is. My contention is that it basically boils down to doing a special kind of harm to them, by forcing them to do more than their proportionate share of labor that by definition is toward a shared end. I will shortly discuss some alternatives to an account anchored on harm, which I think fail, but, toward the initial plausibility of the idea, the very notion that non-voluntary fairness claims can only be generated in the context of presumptive benefits seems to hinge on harm. Klosko primarily writes about the importance of presumptive benefits from the perspective of the freerider, considering possible arguments someone desiring to freeride could conjure up.¹⁶ But, from the perspective of the willing participant, the freerider only matters if the freeriding makes some difference in their life. If we modify the example I gave such that the manual labor required to contain the noxious gas is negligible, and can be discharged in the normal course of a day, the 80 willing and necessary participants should have little reason to care if the remaining 20 residents participate or not. Similarly, if the manual labor actually is costly, but simply requires 80 individuals to be present, and additional bodies are meaningless, then it might be important to have exactly 81 people present, so that each individual is separately free to leave - and so we can imagine, when 81 people are there, each of them having reason to stay, or each person up to the 81st person having a fairness based reason to make the

¹⁵ *Ibid.*, 35

¹⁶ *Principle of Fairness*, 42-43

trip, to alleviate the necessity of each individual's presence. But, beyond that, it makes *no* difference to the contributors whether the remaining residents show up, and neither my reading of the arguments nor intuitions suggest that there should be independent fairness duties.

In short, I think fairness claims are generally about *culpably harmful inaction*. Another compelling fairness example, this time in the mutual security context, is a community defense scenario: invaders are approaching a town, and if they succeed in their attack everyone in the town will suffer grave consequences. However, the town military is technologically sophisticated, and if everyone mans their weapon the invaders will be repelled without significant risk of injury or death. But any lack of participation will lead to defeat. None of the townmembers consented to participate in this sort of collective defense, and an individual sitting out - on claim of pacifism, perhaps - would not be violating any conventional moral principle (though I will suggest in chapter III that such a person would violate a plausible natural duty). I think, again, that Klosko's account is helpful here: the nonparticipant would either freeride on compatriots somehow taking up the slack, or would enable the consequences of defeat, in either case causing harm to his fellow citizens without plausible justification. Among other things, this seems deeply unfair. This sort of analysis will apply to all the best examples for fairness, I think.

Klosko does discuss the role of harm in the enforceability of political obligations, but as a precondition for fairness claims to override liberty claims, rather than as constitutive of fairness claims themselves.¹⁷ This latter acknowledgement suggests an improvement to his discussion of the variable force of political obligations, which on his account depend on the consequences of generalized disobedience of a law (which we are taken to have fairness-based duty to obey).¹⁸ Rather than saying that all cases of disobedience are coequal in terms of violating the principle of fairness, and that the importance of the violation depends on the distinct generalized consequential reasoning, we can simply recognize that significance of the harm (or risk of harm) involved in a case of disobedience directly dictates the strength of the fairness claim, without the intervening heuristic.

This sense of fairness, of course, runs counter to some notable accounts and examples of fairness claims. Justin Tosi gives an intentionally archetypal example of the principle of fairness that might raise questions about harm: a person who hops a turnstile to (literally) freeride on public transport.¹⁹ Tosi correctly points out that this, as discussed previously, should not be interpreted

¹⁷ *Principle of Fairness*, 46-48

¹⁸ *Ibid.*, 99-100

¹⁹ Tosi, Justin. "Playing Fair and Following the Rules." *Journal of Moral Philosophy*, 06, 2016, 5

through tacit consent - by hopping the turnstiles, the freerider quite clearly indicated that they did *not* consent to the system of payment. But, nonetheless, their actions appear wrong. One option that Tosi discards is the argument from ultimate consequences - that the public transit system either *is* harmed by the freeriding, or, as Klosko mentions, that it *would be* harmed if the behavior were generalized. The actual consequence argument is dubious, of course, because of the negligible harm any one freerider does. The generalization argument is tricky in its own right, but Tosi moves past it, and given that I view it as an unnecessary abstraction, I will too. The important contention is that simple *disrespect* for fellow riders is the source of the wrong. My response should be clear enough: assuming that the public transit system is not overcrowded or the like (and the freerider does not somehow cause a breakdown in the system), no appreciable harm is done to the paying riders, so there is no reason for them to feel disrespected. They paid, and the freerider cheated, but they are not meaningfully *paying for* the freerider - another seat on the train just happens to be filled. So why is freeriding in this case, and similar cases, wrong? The simple answer, it seems to me, is that freeriding is an unauthorized use of property, covered under moral principles against *theft* regarding the operator of the public transit system, entirely distinct from the fellow riders. Removing fairness from consideration leaves the moral conclusion unchanged, and in fact greatly simplifies the account.

If various other fairness examples follow this pattern, seeming convincing when referencing culpable harm from inaction (or, less frequently, harm due to genuinely threatening the functioning of cooperative schemes) and unconvincing and likely devolvable to other principles in other cases, it raises a problem for fairness as a grounding for political obligation: fairness is usually in a position to tell people that they *ought* to cooperate in some scheme, which is not a strong position from which to justify a *prima facie* obligation to obey the law (i.e., to *not* disobey the law). Most cases of obeying the law have little if any similarity to the successful fairness examples shown so far, and in a large state the “back door” to fairness claims, where disobedience threatens to take away specific cooperative benefits, is widely acknowledged as implausible. Laws setting speed limits can illustrate this difficulty. A *prima facie* duty to obey the state seemingly should indicate that we have good reason to drive the speed limit all the time, but Klosko argues that, instead, the strength of our duty varies according to how harmful it is - speeding on an empty highway is far different from speeding in a residential neighborhood.²⁰ The argument here is that the disrespect-based account of the basic principle of fairness works to generate a *prima facie* duty, but that this duty is not strong enough to override basic liberty prerogatives unless disobedience is connected to severe

²⁰ *Principle of Fairness*, 99-100

consequences. Both parts of this structure sit oddly with fairness reasoning: my argument has been that fairness duties cannot be generated without internal reference to consequences (harm), but the force of obligations depending on those consequences, I think, is what motivates Simmons' criticism that *all* fairness duties might actually be misplacements of other sorts of natural duties. If fairness considerations cannot ground obedience without the aid of additional consequential reasoning, it begins to appear that the independent duties to not cause whatever set of consequences are really doing the normative work.

The hope for fairness accounts from this point must be to salvage the claim that obedience to general laws pertaining to security is successfully grounded in fairness, and then lean heavily on Klosko's "indirect argument" to justify the many remaining state features. Both steps of that process are questionable: there are remaining questions about how fairness, even if it can justify security provisions in some cases, can justify *modern* levels of security, which lead to doubts about the viability of the indirect argument to justify other features of contemporary states.

One puzzle for fairness which, as far as I know, goes unaddressed in Klosko's account, is the matter of *thresholds* of vitally important goods or benefits. The focus in delineating presumptive goods is on identifying things that people could not live without, but the *level* at which a good is necessary is shortchanged. Security, surely, is presumptively beneficial - but the marginal increase in security provided by more strenuous airport security is not. There is a persistent (and in many cases understandable) equivocation between the provision of a presumptively beneficial good and the *presumptively beneficial* provision of a good. The former is necessary for the latter, but many examples can illustrate their difference. In fact, I think this distinction and the lack thereof is why Simmons' water example mentioned previously is structured the way it is. He also does not distinguish between water being presumptively beneficial - which of course it is - and the town's collective provision of the water being presumptively beneficial - which, due to the individual's independent capacity to dig a well, it clearly is not. This leads him to an ultimately unsatisfying rejection of non-voluntary fairness in general, but the weakness in his objection is perhaps a reflection of the original account.

Once this distinction is brought into play, we need to ask a different set of questions about modern states; namely, merely asserting that states provide security and security is presumptively beneficial is insufficient, even though both statements are true. We can easily imagine a case where a collective institution starts out by providing a presumptive benefit, and justifiably demanding some kinds of cooperation. A group of people band together in a desert to create a purely public water supply - imagine they find a way to increase frequency

of rainfall. Assuming appropriate harm conditions hold (or not, if you reject some of my arguments above), they could demand assistance from everyone in the area. But if they ramp up the scale of their project, and make it rain hard every day, clearly they should not similarly be able to demand cooperation in *that* endeavor - the marginally increased access is not presumptively beneficial anymore, even though the good itself (water) still is. The same structure could be applied to states - sure, security is a presumptive good, but it seems perfectly plausible that lots of people in the contemporary United States might prefer *less* security than we currently enjoy, and thus seemingly should only be obligated by fairness up to the point where additional security stops being presumptively beneficial.

This poses a difficult set of questions for fairness theorists about where presumptive benefits cut off, and how to cope with disobedience in states that go beyond that threshold. More importantly for my purposes, it raises a challenge to Klosko's indirect argument, in which he argues that many modern state features - roads, public health, public education - are inextricably linked to and thus justified by security provision, which is taken as the core presumptive benefit of states.²¹ This conception of discretionary public goods is somewhat conservative on its face - the direct, unavoidable connection between many of these state features and security is unclear to me - but is particularly so in light of this observation. It seems clear to me that most modern states go far beyond the strictly presumptively beneficial in their pursuit of security, and, if we consider the possibility of small, comparatively primitive (but potentially still just) states that focus on providing a minimalistic, libertarian-flavored version of security, it is precisely the features that Klosko identifies as inextricably linked with security that would be left on the cutting room floor. Even if all the rest of the fairness argument goes through, actually expanding that justification through an indirect argument to cover even most of what modern states do, in this light, seems like a difficult and likely impossible task.

QUASI-JUST STATES: A TEST OF NORMS

Now that our attention has been directed to smaller, non-traditional versions of states, let me (much more briefly) raise another kind of puzzle for fairness. Just like small, cooperative, just states could provide security without so many additional provisions, it seems me that less cooperative, arguably *unjust* or at least less just institutions could also provide security, acknowledged as the key good for collective life, without meeting the qualifications to generate fairness obligations for its citizens. What I will suggest is that we can imagine scenarios

²¹ *Principle of Fairness*, 86-95

where there *should* be normative reason to at least minimally cooperate with a state or security provider, but where fairness by its nature cannot power that normative analysis - instead of fairness failing, as I have been focusing on, this is an area where fairness falls short.

An example, to which I will return in subsequent sections: imagine a classic western film premise; a town is overrun by various outlaws, thugs, and gangs. Most people have fled, and those who remain cower in their homes. Suddenly, a nameless cowboy (think Clint Eastwood) arrives, doles out swift punishment to the villains, and order and confidence slowly begin to emerge. This seems like an easily accessible scenario, both in pop culture and the real world; order imposed on chaos through force often, at least initially, looks quite beneficial. In political form, of course this is basically a benevolent dictatorship. The “hero” punishes and gives orders, sometimes wrongly, based on his whims and intuitions. But he is largely benign, and the outlaws, eager for revenge on him and the town, are repelled only by his quick draw and accurate aim. The point of this example is: what is the political duty of the townspeople? Government by lone gunman certainly doesn’t fall under conventional fairness logic - the security benefit that accrues to the townspeople does not come from collective effort or sacrifice, but instead simply falls upon them. Even if one wanted to work around this, and argue this case under some kind of brute fairness logic, whereupon fairness obligations can be grounded in mere receipt of goods that were not cooperative in nature in the first place, Klosko would respond that such an account hasn’t coped with the difficulties in either the distribution of goods in a cooperative or the determination of what exactly counts as “fair.” He’s quite explicit, and understandably so, about requiring not only fair outcomes, but fair procedures in order for institutions to be legitimate bearers of fairness duties.²² But this scenario isn’t clearly just, makes no claims about fair distribution, and is paradigmatically undemocratic.

But if we conclude that there simply is no political obligation here to be found, consider what follows: if one angry person (say, one of only a few people the lone cowboy erroneously harmed) ambushes and shoots the gunman, the outlaws still nearby will flood back into the town, returning it to chaos and violence. The moral harm is more than just the murder; it’s the massive consequences to the polity. To sharpen the point: if a group of townspeople simply *arrested* and jailed the gunman (again, on an accurate accusation of his wrongdoing), the same consequences would be triggered. What ordinarily would be lawful and moral actions actually bring about awful outcomes, in this scenario. Examples involving unlawful government like this one (or a Mafia-run locality) are often used to as counterexamples in accounts of political obligation,

²² *Principle of Fairness*, chapter 3

since duties to despicable regimes seems like something we should avoid. This example is meant to illustrate the opposite: that any account of political duties that *fails* to sometimes justify supporting unjust political systems is insufficient, and in fact quite unintuitive. We should be careful, as well, to note to *whom* it seems like *some* duty should be owed: it isn't that Clint Eastwood now morally deserves obedience through gratitude, or at least it doesn't seem that way to me (the townspeople certainly may owe him debts of gratitude, but the normal objections about that not necessarily be substantiated in obedience apply). The intuitive duty of a person in the town, it seems, is toward the other townspeople. If someone incapacitates Clint Eastwood, in whatever manner, great harm is going to befall *them*.

In chapter 3, I'll elaborate the theory of political obligation that I think falls out from this, and which I argue evades the normal gamut of objections. The point, for now, is this: even if fairness survives or is able to respond to the variety of mechanical objections given (as well as various possible others that I argued against or did not address), and even if the principle does not collapse into other moral principles at key junctures, fairness may not have the ammunition to justify all the political obligations that it should. Note that this is a distinct kind of generality concern; Klosko and others have argued that fairness may not be perfectly general for all the citizens of a state, and that multi-principle theories may be necessary and acceptable to plug any holes.²³ But this is a question of generality across *institutions*, not members thereof - if there are quasi-just states (I'll give some more examples of this in the next chapter) that *should* generate obligations but can't do so under fairness, that speaks to an important gap in the power of the theory, which may be enough to sink it entirely, or should *certainly* be enough, I think, to prefer a different theory which can cover all the states we'd like, while still not tipping over into justifying obedience to any state wielding a sword.

²³ Klosko, George. *Political Obligations*. Oxford: Oxford University Press, 2008

BIBLIOGRAPHY

Klosko, George. *Political Obligations*. Oxford: Oxford University Press, 2008.

Klosko, George. *The Principle of Fairness and Political Obligation*. Lanham, MD: Rowman & Littlefield, 2004.

Simmons, A. John. *Justification and Legitimacy: Essays on Rights and Obligations*. Cambridge: Cambridge University Press, 2001.

Song, Edward. "Acceptance, Fairness, And Political Obligation." *Legal Theory* 18, no. 02 (04, 2012): 209-29.

Tosi, Justin. "Playing Fair and Following the Rules." *Journal of Moral Philosophy*, 06, 2016.

Sexual Consent and
Deception: A Reply
to Dougherty

Ben Atlas
University of Cambridge

I. ESTABLISHING THE PROBLEM

It has long been accepted that the two paradigms of coercion are force and deception—both of which vitiate consent. This paper is concerned with deception: it at first seems uncontroversial to say that deceiving someone into sex is seriously wrong. To illustrate the unethical character of such an act, Dougherty gives two examples. First, if individual A convinces individual B that sex with him is the only way to properly perform a medical procedure, only after which B agrees to sex, it seems A has done something seriously wrong and B has not given valid consent. This type of case can be found in the American court case *Boro v. Superior Court*.¹ Second, if individual A pretends to be individual B's husband, and receives consent to sex under the false assumption that he is B's husband, the encounter seems nonconsensual and A's actions seriously wrong. This type of case can be found in *The Three Musketeers*, when D'Artagnan seduces Milady by pretending to be her husband in a poorly lit boudoir.²

The complexity comes in when we examine what it means to deceive someone into sex. Obviously, not any lie told in the process of gaining consent vitiates consent. If A, in the course of seducing B, claims that Jill Stein of the Green Party received 50 percent of the popular vote in 2016, and B naively believes this to be the case, any consent B gives is not necessarily vitiated by this lie—indeed, the lie is completely unrelated to the sexual encounter. To distinguish between lies that do and do not vitiate consent, Dougherty introduces the concept of a “deal breaker:”

“Second, the deception must concern a deal breaker—a feature of the sexual encounter to which the other person's will is opposed. This requires more than concealing an undesirable feature. It must be the case that the other person is all things considered unwilling to engage in the sexual encounter, given that it has this feature.”³

At first, this seems uncontroversial—lies which counterfactually determine the existence of consent in the first place ought to vitiate consent, and other lies don't seem relevant to the issue at hand. However, when applied, this principle leads to some odd conclusions.

If someone lies about, for example, his or her *alma mater*, this intuitively seems like a lie of the peripheral variety, and not seriously wrong. But such a lie might well be a deal breaker. If person B only wants to have sex with someone who went to Cambridge, and person A, who went to Oxford, lies and claims she

¹ See *Boro v. Superior Court*.

² Dumas, Alexandre. *The Three Musketeers*. Random House, 2011.

³ Dougherty, Tom. “Sex, Lies, and Consent.” *Ethics* 123, no. 4 (2013): 719.

went to Cambridge, B's consent to sex is vitiated—if A had told the truth, B would not have consented, so the encounter is nonconsensual. The conclusion of Dougherty's *Sex, Lies, and Consent* is that lies about one's profession, desire for a relationship, or *alma mater*, which seem only sleazy and minorly wrong, are actually seriously wrong and vitiate consent in the same way that the doctor's lie about sex as a medical procedure does.⁴

To many, this will seem so counterintuitive that something must have gone wrong in Dougherty's analysis. Indeed, Dougherty himself concedes that his argument will "offend some common intuitions,"⁵ but argues that this is simply the cost required for philosophy to challenge conventional wisdom. In my view, placing all lies about deal breakers on a commensurate moral level is an unacceptably high cost to pay, and so we should look for a principle which is more accommodating to our sharply different intuitions about different scenarios. The gap in our intuitions widens when you consider the following categories of exceptional cases.

II. PREFERABLE DECEPTION: A PREVALENT EXCEPTION

Having established the problem, I introduce a category of cases that involve deceptions about deal breakers but sexual encounters that clearly retain their consensual character. Say person A is disgusted by supporters of the Green Party, and would be unwilling to have sex with anyone of that political affiliation. This is because A would be unable to overcome his aversion to a potential partner's politics; knowing of their support for the Green Party would ruin any sexual attraction A felt toward B. Crucially, however, person A would prefer not to know if a potential partner supported the Green Party; indeed, he would prefer to be lied to. In such a scenario, deceiving person A about one's political affiliations does not seem seriously—or perhaps at all—wrong, given that person A actually prefers to be deceived! By Dougherty's lights, however, this would count as lying about a deal breaker. All things considered, if person A knew his sexual partner was a member of the Green Party, he would no longer wish to engage in a sexual relationship with that person. I will call this type of encounter cases of "preferable deception."

The trouble is the way Dougherty applies his concept of a deal breaker. In his example of an irrelevant deception about how much one person likes another's umbrella, he comments that this is not problematic "so long as knowing the truth about the other person's lukewarm opinions of the umbrella would not change the deceived person's willingness to have sex at the crucial

⁴ Ibid.

⁵ Ibid: 721.

moment.”⁶ Even if, after the fact, knowing a sexual partner was a Green Party member would not cause remorse over the encounter, it would change person A’s willingness to have sex at the “crucial moment.” It is evident that knowledge in the crucial moment should not, therefore, characterize what it means to be a deal breaker. An exception must be made for preferred deceptions. A better definition of what it means not to be a deal breaker might be the following:

A deception is not problematic so long as, in a cool moment, knowledge of the deception would not change the deceived person’s willingness to have sex with the deceiver.

It is not immediately clear what is meant by a “cool moment”—it might mean retrospectively, or preemptively, or something else entirely. I do not seek to fully show the details of this here, only to show that Dougherty’s definition of a deal breaker, reproduced below, needs revision:

“It must be the case that the other person is all things considered unwilling to engage in the sexual encounter, given that it has this feature.”

This is because it says nothing about when all things must be considered. If they are considered before the encounter, that consideration might change a person’s willingness to engage in a sexual encounter, and yet still might be such that the person wishes she had not been forced to consider all those facts at all. The needed revision should be something like “all things considered from an objective point of view, removed from the encounter.”

One response open to Dougherty is to argue against the counterfactual interpretation of his deal breaking criteria⁷—if instead it relies on actual mental states, this objection might be avoided. If one’s position is that he is willing to sleep with a Green Party member, but does not wish to know about it, then there is no aspect of the encounter to which his will is opposed, and so consent is not vitiated. I have two responses. First, throughout the paper, counterfactual language is employed—for example the use of terminology such as “crucial moment,” and “all things considered.” So, if nothing else, we may use this category of cases to call for a revision of this language and a rebuttal of a counterfactual account. But second, and perhaps more importantly, I believe that there is no way around some counterfactual element within the definition of a deal breaker. It is simply not the case that people come into situations with a clearly laid out set of intentions that applies to every possible encounter. Rather, the way people determine their intentions is inherently counterfactual. It might be that I do not have an explicit principle against sleeping with Green Party

⁶ Ibid: 719.

⁷ This objection was raised in informal correspondence with Dr. Dougherty. Even if we take it as sufficient for avoiding this category of case, revision to much of the counterfactual language in *Sex, Lies, and Consent* is needed.

members formulated in my mind—it is only when this possibility is attended to that my will shows itself to be so opposed. Determining my will forces me to consider something counterfactual—if she were a Green Party member, would that change my decision? The question remains: when should this be considered? In my view, it must be from some removed viewpoint, and importantly not at the “crucial moment” before sex.

My final point under this objection is that encounters which would count as rape-by-deception using Dougherty’s criteria are usually, or at least often, cases of preferable deception. This is an empirical point, and not one that I have the data to corroborate, but intuitively true nonetheless. It seems like most, though not all, instances of rape-by-deception occur in casual encounters outside of relationships (it would be unusual, for example, to lie to a relationship partner about one’s *alma mater* or political affiliation throughout a relationship). Indeed, Onora O’Neill views casual sexual encounters as immoral⁸ precisely because the risk of deception is so high—after all, there is probably *something* about you which, if considered at the crucial moment, would change a potential sexual partner’s willingness. In Dougherty’s own example, of a soldier and a pacifist who engage in a one-time sexual encounter, it seems likely to me that the pacifist would in fact rather not know about the soldier’s past violent actions—after all, is it not irrational to differentiate between one-time sexual partners based on something wholly irrelevant to the act itself? Isn’t what really matters to the sexual encounter, and therefore the parties to it, the *belief* that you are having sex with a pacifist? Such preferable deception is not always the case, of course, but I hypothesize that it applies to a large number of casual sexual encounters. This would partially explain why people usually do not feel violated in the same way after learning their partner did not, in fact, go to Cambridge, as they do after finding out they were not, in fact, having sex with their husbands—the belief that one’s sexual partner went to Cambridge is what they are after, as opposed to the objective truth that their sexual partner is their husband.

Furthermore, it is an absolutely unacceptable price to pay if we lump this category of cases in with the medical practice deception and the D’Artagnan deception. As Sherry Colb notes,⁹ people routinely wear make-up, deodorant, and other cosmetic enhancements—if a potential sexual partner knew what one *really* looked and smelled like without these, it may well change their mind in the crucial moment. But, it would “utterly trivialize the crime of rape to suggest that wearing hair plugs or using deodorant might convert consensual sex into rape.”¹⁰ The key to avoid this is that the other sexual partner would rather *not*, at the crucial moment, have the unpleasant knowledge of what someone smelled

⁸ O’Neill, Onora. “Between consenting adults.” *Philosophy & Public Affairs* (1985): 252-277.

⁹ Colb, S. F. “Rape by Deception, Rape by Impersonation, and a New California Bill.” *Verdict/Justica* (2013).

¹⁰ *Ibid.*

and looked like without deceptive cosmetic enhancements; and so, nothing like rape-by-deception has been committed. If we count preferable deception cases as seriously wrong, we are committed to the widespread condemnation of clearly consensual, harmless, innocent sex—and the subsequent trivialization of actual rape.

III. DEAL BREAKING UNDER THE CIRCUMSTANCES OR IN A VACUUM?

Next, I introduce a second category of exceptions to Dougherty's criteria: instances in which one is deceived into a sexual encounter which they deem desirable and pleasurable generally, but would rather sleep with someone else. To illustrate this objection, consider the following example: Person A is attracted to person B, and thinks it would be enjoyable to have sex with B. However, Person A only wants to have sex with one person tonight, and therefore would prefer to have sex with the most attractive person in the room. Unbeknownst to Person A, Person B's friend is a model, and far more attractive than Person B. What's more, B knows that A would find his friend more attractive than himself. A asks B: "I want to have sex with the most attractive person here—is that you?" B replies: "Yes!" and the two spend an enjoyable night together. This, of course, meets Dougherty's criteria for rape-by-deception; Person B has lied about something which, had Person A known the truth of, would've changed Person A's willingness to have sex with Person B. I will call such cases "generally enjoyable, relatively undesirable" deceptions.

There is a strong intuition that Person B has not done anything seriously wrong in this instance—after all, both A and B ended up having a good time, with no substantial regrets. Even if A knew of B's lie *post facto*, her reaction would almost certainly not be comparable to the reaction of a patient deceived into sex as a medical procedure. Dougherty might say that this is due to false intuitions about lying and consent, but I believe there is a deeper difference. Both are nonconsensual—there is no denying that—but they are nonconsensual in different ways, and in each case the subject of consent is fundamentally different. I pursue this line of thought further in Section V while trying to reconcile the cases mentioned with a revised principle of consent.

IV. IMMORAL NON-CONSENT

The third and most controversial category of exceptional cases concern instances of people acting immorally in their refusal to consent. If someone is deceived into an action which, all things considered, they would consent to if they were acting morally, has their consent been vitiated in a morally

problematic way? Consider the following case. Person A is an African-American, but with an ethnically ambiguous appearance, chatting with Person B. Both make it clear they want to have one-time sex with one another without any further commitment. As they walk up the stairs toward Person B's apartment, Person B remarks to Person A: "Wait, you're not black, are you? I don't want to have sex with a black person." Person A considers the situation, and figures it will be harmless to answer in the negative. Both have an enjoyable evening and never see each other again. Can we really say that Person A has done something unethical to Person B?

Perhaps we can; after all, Person B would certainly feel violated if she learned of Person A's ethnicity *post facto*. But I suspect it will not sit easily with most readers to assert that this case deserves the same sanction as one in which the aggrieved party's refusal consent was *not* immoral. For example, most thinkers (except some particularly committed libertarians) view laws against businesses discriminating based on race, gender, or sexual orientation as perfectly acceptable. The government conscripting a barber to provide haircuts for the military does not seem morally on par with forcing a barber to provide his services equally to gay and straight citizens. I posit that one's consent becomes less morally considerable when it is based on obviously unethical considerations.

V. TOWARD A NEW THEORY OF SEXUAL CONSENT

These three categories of cases evidently call for a revision of Dougherty's thesis. I believe each presents a distinct problem, and therefore requires distinct revisionary analysis. I have already showed how we might fit the first category, preferable deception, into Dougherty's definition—consent is only vitiated if some aspect of the encounter is lied about and, all things considered from a point of view removed from the encounter, that lie would cause the aggrieved party to withdraw his or her consent. There are, as I mentioned, difficulties with the concept of such a removed point of view; indeed, Bernard Williams once wrote in his criticism of Henry Sidgwick that ethical theorizing from such a standpoint is "not very sophisticated,"¹¹ and questioned the objectivity of its stance.¹² But, I do not think it is acceptable to attend to this category of cases without some revision to Dougherty's thesis—to "bite the bullet," so to speak—as that assigns serious blame to wholly innocent actions, trivializing cases of actual rape-by-deception. Perhaps there is a better, more elegant way to revise

¹¹ Williams, Bernard. "The point of view of the universe: Sidgwick and the ambitions of ethics." *Cambridge Review* 103 (1982): 167.

¹² *Ibid.*

the criteria without recourse to a “cool moment,” but some revision must be made to account for this prevalent category of cases.

The second category, of broadly desirable but technically nonconsensual actions, reveals something deeper about the nature of sexual consent. In my view, there is no need to revise Dougherty’s criteria at all to account for this category, only the conclusions we can draw from it. Dougherty characterizes sexual consent as binary—consensual sex is morally acceptable, and nonconsensual sex is seriously wrong. The trouble with this formulation is that not all instances of nonconsensual sex are equivalently bad. Having sex with a random doctor who is supposedly performing an abortion seems extremely undesirable, but having sex with someone who you genuinely enjoy having sex with, but would not be your first choice, seems not so bad. So while both cases are nonconsensual, the difference might be akin to that between purposefully stepping on another’s foot and punching him in the face. The object of consent has crucially changed, and so has the moral status of each action. It is perhaps minorly wrong to misrepresent oneself as the most desirable mate in the room, but only insofar as it has non-consensually decreased a potential mate’s satisfaction by a comparatively small amount. This is an important consideration, not just for cases of sexual consent, but consent in general—it is not just the existence of consent which determines an action’s moral status, but also the subject of consent.

The third category, regarding immoral refusal to consent, is the most difficult to incorporate in a theory of sexual consent. This is because, unlike the other cases, it involves a clearly nonconsensual action that the subject of which finds extremely undesirable. One way to circumvent this would be to bring in some objective standard for good and bad—for example, since it doesn’t really matter what race a sexual partner is, we could say that withdrawing consent due to something wholly irrelevant to the act is irrational, and can so be overridden without serious moral consequences. However, this goes too far. D’Artagnan could argue that whether or not he is married to Milady is irrelevant to the actual sex, and so as long as she never finds out, he has done nothing wrong. Another approach might be to simply protect this class of cases by adding a provision that discriminatory or otherwise unethical criteria for consent are not legitimate—but this is problematic in its objective prescription of what is and is not ethically legitimate deal breaking criteria. After all, what matters is not whether or not some enlightened community of philosophers thinks your consent criteria is valid, but rather whether or not *you* think your consent criteria is ethically valid. Finally, we might try to use the same approach as on the second category of cases. After all, it seems like the subject of nonconsensual sex who unwillingly has sex with an African-American has been tricked, but into something which is not bad at all. But this runs into the same problem—who is the authority on what

is actually good for someone, if not the person themselves? I conclude that this category of cases involves actions that are, despite my initial intuition, seriously wrong. Perhaps the muddled intuitions here come from the fact that both parties have done something wrong—as it is unethical to discriminate between sexual partners based purely on ethnicity—and so an element of *schadenfreude* enters the analysis. But this hardly seems like legitimate philosophical analysis, and I cannot think of any way to account for this category of cases without asserting unacceptable paternalism.

VI. CONCLUSIONS ABOUT DIFFERENT CASES

Finally, I want to reexamine the intuitions, which Dougherty purports to overturn, in light of my previous analysis. I believe that much of his criticism, in the real world, tends not to apply. This is because, as previously noted, most cases in which people only want to have sex with someone from Harvard are actually cases in which they are fine having sex with someone who they think is from Harvard, or at the very least not seriously opposed to having sex with someone who is not from Harvard. I hypothesize (without actual empirical data to support my claim) that most real-world cases fit the categories of preferable deception or involve consenting to something that is only relatively, rather than wholly, undesirable—and in many cases, both. For the most part, cases in which people are deceived into sex are not comparable to D'Artagnan's deception of Milady, or the doctor's deception of a patient. Nevertheless, this scrutiny of sex, lies, and consent yields some surprising conclusions. Deception into sex in the third category of case, immoral non-consent, seems to involve something seriously wrong. Furthermore, some cases of rape-by-deception do not fall into one of these categories: for example, cases where one partner lies about wanting a relationship, or about their religious affiliation, or something similar. It is also *possible* that something like alma mater would be a genuine deal breaker; it just seems practically unlikely. While this hinges on empirical likelihood of cases fitting into the first or second category of exception, I also believe these mitigating circumstances are enough to rebut the claim that casual sexual encounters come with a high degree of moral hazard. The exceptions raised serve not only to point out gaps in Dougherty's narrative, but also provide a less counterintuitive alternative, which avoids placing serious blame on actions which are not seriously objectionable.

Husserl's 'Science of Pure Consciousness'

Daire Boyle
Maynooth University

The phrase ‘has science gone too far?’ has long been reduced to a lazy attempt at humour by those attempting to react to incredible scientific breakthroughs shattering their previously held conceptions of reality. However, there was a time when this seemingly naïve sensationalism actually held some meaning, particularly for one man, and wasn’t just an irrelevant cliché. We speak, of course, of the great German philosopher Edmund Husserl, who helped lay the foundations for the branch of philosophy concerned with the conscious self – phenomenology. The treatment of consciousness in philosophy, psychology and particularly science from Galileo’s time to his day left Husserl uneasy. He believed that the natural sciences (*naturwissenschaften*) were not adequate to explain everything in our world and that a serious analysis of the subject experiencing said world needed to be carried out in order to give a basis to what he called the ‘spiritual sciences (*geisteswissenschaften*)’ or humanities. This phenomenology formulated by Husserl was supposed to serve as a means to analyse that which we experience away from the narrow scope of empiricism and psychological assumptions about human nature.¹ This essay shall guide you through Husserl’s criticism of certain elements of the aforementioned natural sciences and shall examine his claim that a ‘science of pure consciousness’ must be a foundation stone for scientific inquiry. We shall argue that Husserl’s critique of the treatment of science of his day was indeed valid and necessary, while also uncovering some of the holes in the formulation of his phenomenological methods and postulations.

We begin our examination of Husserl’s critique of the idealization of nature scientifically by focusing our attention on his responses to the works of Galileo and Descartes. His views on science’s place in philosophy and its usefulness as a tool for absolute objectivity were crystallized in his seminal work *The Crisis of European Science and Transcendental Phenomenology* (1936). This tome was, and indeed still is, quite provocative in its offerings as Husserl sought to get at the heart of the ‘Galilean’ – as he categorised it – scientific revolution. He took issue with the constant striving for a ‘mathematization of nature’² and wished to intellectually reconstruct ‘the train of thought that motivated it’. The paradigm shift that Galileo brought to modern science, namely the formalization of nature as an abstracted mathematical grid, is the main idea that Husserl focused his energies on. Indeed, he recognised that the Galilean conception of science consisted of a *geschlossenheit*, or ‘abstracted closure’, whereupon the lived experience of phenomena in nature were secondary to this aforementioned mathematical grid. He went on to argue that mathematical science enforced this grid on all things and as a result covered all it examined in an *ideenkleid*, or ‘garb of ideas’, abstracting away the ‘real world’ that we experience. This

¹ Dermot Moran, *Introduction to Phenomenology*, (London: Routledge, 2000), p.5.

² Dermot Moran, *Husserl’s Crisis of the European Sciences and Transcendental Phenomenology*, (New York: Cambridge University Press, 2012), p.66.

recognition of an abstraction from the tangible and experiential echoes the earlier works of Descartes, Kant and Hume, who posited similar ideas relating to this dualism of the self in the world. Husserl credits Galileo with being the first to expound the notion of ‘nature as a self-enclosed world of bodies’.³

Perhaps the salient point to take away from Husserl’s analysis of Galileo is that he does not deny that ‘Galileo’s world’ of the abstract mathematical objects isn’t real, rather that the sense in which it is real is not fully grasped by Galileo.⁴ He rejects the notion that mathematical science is a tool to uncover the “true” world as it exists “in itself”.⁵ While Descartes definitely influenced Husserl’s thoughts regarding the abstraction of nature, Husserl stops short of fully endorsing all the ‘main’ Cartesian principles. In particular, Husserl takes umbrage with the Cartesian teachings on dualism, which separated the ego from the rest of the world as it was ‘supposed to be’. For Descartes the ego was the mechanism that carried out thinking but was separate from the body and even from thoughts themselves. This was simply unedifying to Husserl; to his mind Descartes ‘puts out of action, with one blow’ all knowledge of the world including those of the straightforward experience.⁶ Descartes’ whole approach was indicative of the overall problem in the methodology of science in Husserl’s mind. He believed that Descartes adopted such radical doubt in order to underline his pre-determined position that empirical science was the only legitimate and objective way of achieving truth.⁷ We can easily see the rationale behind Husserl’s position at this point in time; when one assumes or proposes an outcome, it can be quite easy to then construct a scenario or framework within which that outcome occurs. Furthermore, Husserl’s assertion that one cannot separate the mind and body like Descartes suggests seems well founded; reality is surely applicable to all parts of man and not simply the ego. Everything must be cloaked in the world of sensory experience. One can find the influence of this dogma in the writings of Maurice Merleau-Ponty, who espoused the notion of the philosophical ‘flesh’ between the self and the sensory experience. This idea of the self not being a *barrier* to examination of the outside, but rather as an *essential part* of that examination is one that can be nicely conceptualised and now, although at the time this theory was quite the departure.

It is here that we see that Husserl’s real issue was with the logical positivists of his day, notably the Vienna Circle led by Moritz Schlick. The Circle

³ Ibid., p.69.

⁴ Gail Soffer, ‘Phenomenology and Scientific Realism: Husserl’s Critique of Galileo’, *The Review of Metaphysics*, 44.1 (1990), 67-94 (p. 69)

⁵ Moran, p.70

⁶ Edmund Husserl (translated by David Carr), *The Crisis of European Sciences and Transcendental Phenomenology*, (Northwestern University Press, 1970), p.170.

⁷ Ibid, p.173.

were heavily influenced by the teachings of Auguste Comte and Ernst Mach and were concerned with the *meaningfulness* of statements. They set out about eradicating metaphysics by adopting a logical approach to all philosophical problems as propounded by classical empiricists such as Hume and Locke. Members of the Circle judged whether a statement had meaning by considering its constituent words and message – if its terms were logically consistent and its embedded message was verifiable then the statement was said to have sense, otherwise it was a nonsensical statement. In Husserl's view, the positivists had 'decapitated'⁸ philosophy as their ruthlessly logical approach eliminated the contribution of subjectivity. This subjectivity was necessarily purged by the positivist dictum of science being objective, inductive and experimental which ran contrary to Husserl's views of the nature of science as the formation of an 'all-consuming methodology'.⁹ Indeed this obsession over data collected rather than the process of collection did not sit well with Husserl as it completely missed the centrality of the human experience and consciousness in said activities.

We shall inspect one further group of people, the neo-Kantians, whose philosophy was at odds with Husserl's before assessing his polemic for the 'science of pure consciousness'. The 'arid intellectualism'¹⁰ presented by the followers of the great German philosopher Immanuel Kant's beliefs manifested itself in their fixation of establishing a basis for the *a priori* application of mathematics to nature.¹¹ This objective followed on from Kant's maxim that Newtonian science should be taken as 'fact'.¹² This conceptualisation of science is certainly an intuitive one and one that bears out under examination. However it is hard to shake the notion that perhaps Kant's phenomenology was too rigid in its scope, too clumsy an implement to truly encapsulate the unique being and position of the observer who ascertained the *facticity* of that framework to begin with. More precisely, Kant proposed that the Newtonian world view and its notions of space and time were directly attainable through *a priori* human sensory interaction. Here Husserl steps in to lay out the idea that perhaps this science is a formal construction and is not necessarily isomorphic to lived, human experience.¹³ He urges us to pause and scrutinise whether or not the previously alluded to world of the abstract mathematical grid is indeed the *real* and incontrovertible world we live in; corresponding directly to the experiential one with which we interact.

It is against this backdrop that we can now investigate Husserl's claim

⁸ Moran, p. 71.

⁹ Ibid.

¹⁰ Richard Kearney, *Modern Movements in European Philosophy: Phenomenology, Critical Theory, Structuralism*, (Manchester: Manchester University Press, 1994), p.14.

¹¹ Moran, p.71.

¹² Ibid.

¹³ Ibid.

that a 'science of pure consciousness' is a necessary presupposition for scientific inquiry. This claim is underpinned by the fact that at the time consciousness was treated as an entity among other things, rather than any sort of 'spiritual' relationship between objects¹⁴ and that empirical psychology seemed to treat consciousness as mere 'epi-phenomenon'. As a result then, according to Husserl, this view of consciousness necessarily precludes any concrete and 'presuppositionless' basis for knowledge being formed. Flipping this view on its head he argued that rather consciousness itself should be the starting point for inquiry and, accordingly, the rigorous science he set out to germinate was the idea of phenomenology. While Kant, Hume and Brentano all alluded to elements of phenomenology in their own works none quite formalised the field quite like Husserl. It is a method, rather than an idea, with the purpose of getting to the real truth of matters and describing *phenomena* 'in the broadest sense of whatever appears [...] as it manifests itself to consciousness'.¹⁵ This was indeed a novel practice of philosophy and perhaps dealt with ideas of the self, ego and consciousness more satisfactorily than the field of metaphysics and even the teachings of Descartes, allowing for a framework which could overcome the issue of the mind-body dualism.

Husserl states that all that is experienced by an observer is legitimate – within reason: '[e]very type of firsthand intuiting forms a legitimate source of authority'.¹⁶ While this is again one of the more intuitive aspects of Husserl's teachings, it does not hold up as well under further scrutiny. In particular, when a generalisation or assertion is used to extend a set of characteristics to the *genus* of humans as a whole we must be careful; is it the case that that which we experience is legitimate? What of the case whereby the observer is deceived in some way, is the authenticity of his experiences compromised in some way, and if so, what then for the phenomenological outlook as a whole? He eventually ends up at the conclusion that it is human consciousness, rather than the thing of outer sense perception, which has absolute existence. This, then, is the cornerstone upon which all of science should be founded. As discussed earlier there is to be a clear delineation of the different worlds we *describe* and exist in according to Husserl. Hence the abstract world of the mathematical object must be built on top of the clear *lived* world. Describing this world Husserl remarked 'I am aware of a world [...] I discover it immediately, intuitively, I experience it'.¹⁷ This must be the beginning of all science as it the world we actually live and experience, certainly we can apply the mathematical structures to it, says Husserl, however these structures can never replace or degrade the lived, natural

¹⁴ Kearney, p.13.

¹⁵ Husserl, p.20.

¹⁶ D. Sinha, 'Phenomenology and Positivism', *Philosophy and Phenomenological Research*, 23.4 (1963), 562-577 (p. 562).

¹⁷ Edmund Husserl, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy: Book 1*, (Kluwer Academic Publishers, 1989), p.101.

world they are built upon. To my mind this seems like a perfectly acceptable and rational way of viewing phenomena; we cannot divorce the lens of our experience from that which we experience, therefore why should that lens not factor into our descriptions of the world? The mathematization of nature does indeed seem to lend itself to a scorched earth view of the world whereupon our lived experiences simply do not feature as entities. Yet something impels these experiences, something gives shape and colour to the phenomena that bombard us; surely this consciousness should be recognised as the legitimate starting place of our scientific inquiry.

In order to more fully analyse this world Husserl introduced the method of phenomenological reduction or *époché*, akin to Descartes' 'radical doubt' which served to strip back the natural world to its supporting structure. This method relied heavily on the idea of 'bracketing', where all the presuppositions of an object are laid aside and it is examined solely on the attributes it presents to an observer experiencing it. For example, examining a feather would require the spectator to suspend the natural beliefs that feathers are parts of birds, used for flying, are soft and so on. Rather the spectator would simply view the feather as a number of structures that appear in certain ways to their consciousness. Martin Heidegger perhaps presents the most important critique of Husserl's phenomenology and in particular takes issue with the notion of bracketing, stating that bracketing 'of the factual world [...] must be a crucial mistake'.¹⁸ Heidegger's approach to the reductions therefore differed to Husserl's; he put forward the idea of the phenomenological reduction of a being as leading one to 'the understanding of the being of this being' rather than to Husserl's 'transcendental life of consciousness'. This ontological viewpoint does intuitively seem more palatable; when we strip back away the natural attitude from a person's interaction with a phenomenon it surely leaves us with a picture of why the interaction took place in the way that it did. To reference the lens of experience again; it has been applied, once we strip it away we learn more about how that lens was formed in the first place, how that observer came to perceive things as they do. This 'ontological difference' is the main distinction between Husserl's and Heidegger's work on phenomenology, and Heidegger devoted much of his writings to the description of the *essence* of being rather than temporary nature of existing. In this way he argued that there is no pure conscious self, rather we are observers *thrown* into experiential situations. This directly clashes with Husserl's notion of the 'transcendental life of consciousness' and it seems to be one of the few holes in his phenomenological thesis, albeit a major at that. Perhaps Husserl's phenomenology is not fully-formed in its own right and needs the structure of something akin to Heidegger's *Dasein*

¹⁸ D.E. White, *Phenomenology*, (The Internet Encyclopedia of Philosophy: 2015), <www.iep.utm.edu/phenom/> [accessed: 7th May 2015].

underlying it. One could then make the case that the notable lack of the 'pure I' in the Husserlian framework could be avoided by recognising the *mode* of the observer at a particular point in time, by examining not only how they interact with the world, but also by reflecting upon the *impact* these experiences have then on the observer.

In conclusion I find myself drawn to Husserl's unique and thought-provoking critique of Galilean science. He is compelling in his arguments for this science being in crisis due to it having lost its 'sense of rootedness in man's life-experience'¹⁹ and presents a nuanced and fresh way for science to sublimate its focus through the lens of pure consciousness. It does seem that this 'science of pure consciousness' he proposes isn't perhaps fully rigorous and complete, as Heidegger's refinements show us, but overall I agree with the thrust of his position. It seems self-evident that science should stem from the point where it is formulated, namely the consciousness of its adherents. Husserl showed that consciousness simply cannot be discounted when gathering data and formulating theories about the real world we experience around us. I find myself agreeing with his conclusion that science should be about establishing a clear and coherent methodology which takes into account its human conduit rather than being purely about collecting data and I feel that in such a world view a 'science of pure consciousness' is indeed a necessary presupposition for scientific inquiry.

¹⁹ Kearney, p.13.

BIBLIOGRAPHY

- Beyer, Christian, *Edmund Husserl* (The Stanford Encyclopaedia of Philosophy, Edward N. Zalta (ed.), Summer 2015 Edition), <<http://plato.stanford.edu/archives/sum2015/entries/husserl/>>.
- Gurmin, Dr J. Haydn, *Lecture Notes: Phenomenology*.
- Holmes, Kceniya, *What is the significance of the mathematization of nature?*, <http://www.academia.edu/3527765/What_is_the_significance_of_the_mathematization_of_nature>.
- Husserl, Edmund, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy* (Kluwer Academic Publishers, 1989).
- Husserl, Edmund, *Introduction to Logic and Theory of Knowledge* (Springer Publishers, 2008).
- Husserl, Edmund (translated by David Carr), *The Crisis of European Sciences and Transcendental Phenomenology*, (Northwestern University Press, 1970).
- Kearney, Richard, *Modern Movements in European Philosophy: Phenomenology, Critical Theory, Structuralism* (Manchester: Manchester University Press, 1994).
- Laporte, Jean-Marc, 'Husserl's Critique of Descartes' in *Philosophy and Phenomenological Research*, 23 (1963), 335-352.
- Mohanty, J.N. and McKenna, William R., *Husserl's Phenomenology: A Textbook* (Washington D.C.: Center for Advanced Research in Phenomenology & University Press of America, 1989).
- Moran, Dermot, *Introduction to Phenomenology* (London: Routledge Press, 2000).
- Moran, Dermot, *Husserl's Crisis of the European Sciences and Transcendental Phenomenology* (New York: Cambridge University Press, 2012).
- Sawicki Marianne, *Edmund Husserl* (The Internet Encyclopedia of Philosophy, 2015) <<http://www.iep.utm.edu/husserl/>>.
- Sinha, D., 'Phenomenology and Positivism' in *Philosophy and Phenomenological Research*, 23 (1963), 562-577.
- Smith, David Woodruff, *Phenomenology* (The Stanford Encyclopaedia of Philosophy, Edward N. Zalta (ed.), Winter 2013 Edition) <<http://plato.stanford.edu/archives/win2013/entries/phenomenology/>>.

Soffer, Gail, 'Phenomenology and Scientific Realism: Husserl's Critique of Galileo' in *The Review of Metaphysics*, 44 (1990), 67-94.

Weyl, Hermann, *Philosophy of Mathematics and Natural Science* (Princeton: Princeton University Press, 1949).

White, D.E., *Phenomenology* (The Internet Encyclopedia of Philosophy, 2015) <www.iep.utm.edu/phenom/>.

