

School of
Law



BRITISH JOURNAL OF AMERICAN LEGAL STUDIES

VOLUME 3 - ISSUE 1
Spring 2014

ARTICLES

Lost in Translation: Criminal Jury Trials in the United States

Michael G Heyman

Rawls, Political Liberalism, and the Family: A Reply to Matthew B O'Brien

Greg Walker

The Origination Clause: Meaning, Precedent, and Theory from the 12th to 21st Century

Priscilla HM Zotti & Nicholas M Schmitz

From Lark Rise to The Storied City

John Martinez

Commanding Consistently With Sovereignty: Thomas Hobbes's Natural Law Theory of Morality and Civil Law

Kody W Cooper

Melting Pot Benevolence and Liberty Patriotism: The Importance of the Moral Cosmopolitanism Precedent

Will Sarvis

The Expanding Right to an Effective Remedy: Common Developments at the Human Rights Committee and the Inter-American Court

Valeska David

Editor-in-Chief: Dr Anne Richardson Oakes, *Birmingham City University*

Associate Editors

Sarah Cooper, *Birmingham City University*
Dr Haydn Davies, *Birmingham City University*
Prof Julian Killingley, *Birmingham City University*
Dr Jon Yorke, *Birmingham City University*
Seth Barrett Tillman, *National University of
Ireland, Maynooth*

**Birmingham City University Student
Editorial Assistants 2014-2015**

Natika Richards-Perry	Joshua Longhorne
Nicketta Browne-Marke	Ghalid Mohamed
Lee Glass	Susan Nasser
Ellis Isherwood	Amna Nazir

Graduate Editorial Assistants 2014-2015

Daniel Gough	Gareth Lee
--------------	------------

**Maurice A Deane School of Law, Hofstra
University New York Student Editorial
Assistants 2013-2014**

Dave Ackerman	Michael Smith
Pooja Bhutani	Adam Theo
Andrew Garbacz	Kelsey Walker
Omar Harding	

**Maurer School of Law, Indiana University
Student Editorial Assistants 2013-14**

Ron Blue	Jessica Fronk
Sarah Domin	Adesuwa Ighile

Editorial Board

Hon. Joseph A Greenaway Jr., *Circuit Judge 3rd Circuit, U.S. Court of Appeals*
Hon. Raymond J McKoski, *Circuit Judge (retired), 19th Judicial Circuit Court, IL. Adjunct
Professor of Law, The John Marshall Law School, Chicago, IL*
Prof Antonio Aunion, *University of Castille-la Mancha*
Prof Francine Banner, *Phoenix School of Law, AZ*
Prof Devon W Carbado, *UCLA, CA*
Dr Damian Carney, *University of Portsmouth, UK*
Dr Simon Cooper, *Reader in Property Law, Oxford Brookes University*
Prof Randall T Coyne, *Frank Elkouri and Edna Asper Elkouri Professor, College of Law,
University of Oklahoma, Norman OK*
Mark George QC, *Garden Court Chambers, Manchester, UK*
Prof Larry Hammond, *Osborne Maledon PA, Phoenix AZ*
Dr Carolyn Hoyle, *Reader in Criminology, University of Oxford, UK*
Prof James Kousouros, *CALS Visiting Professor, NY*
Prof Ian Loveland, *City University London, UK*
Prof James Maxeiner, *Center for International & Comparative Law, University of Baltimore
School of Law, Baltimore, MD*
Prof Ruth Miller, *University of Massachusetts, Boston, MA*
Dr Anne-Marie O'Connell, *University Toulouse 1 Capitole*
Prof Austin Sarat, *William Nelson Cromwell Professor of Jurisprudence & Political Science,
Amherst College, MA*
Dr Stephen W Smith, *University of Birmingham, UK*
Prof Carrie Sperling, *Sandra Day O'Connor College of Law, Arizona State University, AZ*
Prof Clive Stafford Smith, *Reprieve*
Dr Timothy Stanley, *Rothermere American Institute, University of Oxford, UK*
Prof Adam N Steinman, *Seton Hall University School of Law, Newark, NJ*
Prof Russell Wheeler, *The Brookings Institution, Washington DC*

**BRITISH JOURNAL OF
AMERICAN LEGAL STUDIES**

VOLUME 3 - ISSUE 1

Spring 2014

ARTICLES

- Lost in Translation: Criminal Jury Trials in the United States
Michael G Heyman 1
- Rawls, Political Liberalism, and the Family:
A Reply to Matthew B. O'Brien
Greg Walker 37
- The Origination Clause:
Meaning, Precedent, and Theory from the 12th to 21st Century
Priscilla HM Zotti & Nicholas M Schmitz 71
- From *Lark Rise to The Storied City*
John Martinez 141
- Commanding Consistently With Sovereignty: Thomas Hobbes's
Natural Law Theory of Morality and Civil Law
Kody W Cooper 165
- Melting Pot Benevolence and Liberty Patriotism: The Importance
of the Moral Cosmopolitanism Precedent
Will Sarvis 197
- The Expanding Right to an Effective Remedy: Common Developments
at the Human Rights Committee and the Inter-American Court
Valeska David 259

LOST IN TRANSLATION: CRIMINAL JURY TRIALS IN THE UNITED STATES

Michael G. Heyman*

The John Marshall Law School

ABSTRACT

Having taught Criminal Law for several decades, I've watched students struggle to grasp its basic concepts and language. Over time, we've managed to master these matters. But criminal trials require untrained juries, assembled from the general population, not only to listen carefully throughout these trials, but then to apply an entirely new body of concepts, effectively making decisions about personal liberty and even life itself. Miscues are clearly predictable.

This Article examines the causes of these miscues, while avoiding the simplistic solution of ridding the system of jury trials. Rather, the machinery of criminal justice is flawed by the presence of arcane, outdated criminal codes, judges passively presiding over trials by rarely attempting to prevent or eliminate confusion among jurors and a system of judicial review that presumes correctness and understanding in the face of flatly contradictory facts. Finally, as the title of this piece suggests, much can be improved and justice more fully served by simply reworking the judge-jury relationship, whereby judges more actively and sensitively educate juries on their functions and the applicable law throughout these trials.

CONTENTS

I.	INTRODUCTION.....	2
II.	COMPREHENSION ISSUES.....	4
III.	NULLIFICATION AND OTHER DISTORTIONS OF THE LAW.....	7
	A. JURY NULLIFICATION AS JURY ERROR.....	10
IV.	EDUCATION JURIES.....	14
V.	EDUCATING JUDGES.....	19
	A. EDUCATING JURIES BETTER.....	21
	B. THE TRAYVON MARTIN CASE.....	24
VI.	THE PROBLEM WITH BAD CRIMINAL CODES.....	29

A. LOUISE WOODWARD: A BAD NANNY, BUT NO

MURDERER 31
VII. CONCLUSION..... 34

I. INTRODUCTION

Imagine you have been asked to explain the American criminal justice system to a group of foreign lawyers. Imagine further that this audience easily understands our basic legal structures, including our police system and legal infrastructures. At that point, you venture to explain our jury system, which though hardly unique, can look remarkable from a detached perspective.¹

The perplexity would begin as you describe the civic responsibility for jury service, and that jurors undergo substantial questioning before the jury is finally empanelled.² Further, you explain that, without any additional training or required understanding of law, juries are to render verdicts in criminal cases. Moreover, this takes place after they sit on a case having received nearly no explanation of what was to follow during the trial. Indeed, only at the conclusion of the trial will the judge read the instructions, and she does not invariably provide a written copy of them or any additional explanatory material.

Worse, if you are attempting to paint in broad strokes about the American system, you have to explain the variations in the law nationwide in all their complexity. Thus, though the Model Penal Code prompted an enormous reform movement,³ its simplicity has eluded the codes of many states. Thus, for example, whereas the Code reduced the welter of mental states to four, many states either retain common law notions through their codes, or do not provide a code basis for them at all.⁴ Some leave it entirely

*Michael G. Heyman, Professor, The John Marshall Law School, 315 South Plymouth Court, Chicago, IL 60604.

¹ The United States stands as one of a few Western nations to retain the death penalty, and juries play a key role there. See *Abolitionist and Retentionist Countries*, AMNESTY INT’L, <http://www.amnesty.org/en/death-penalty/abolitionist-and-retentionist-countries> (last visited July 31, 2013). Moreover, ours is also a constitutionally driven reliance on juries for crimes of sufficient magnitude.

² Jurors will be questioned on a variety of items during *voir dire* to determine any potential bias or prejudice.

³ MODEL PENAL CODE Pt. I: General Provisions (Official Draft and Revised Comments 1985), Pt. II: Definition of Specific Crimes (Official Draft and Revised Comments 1980). The Code, officially promulgated in 1962, provides the basis for the codes in the dozens of states that have adopted it in some way.

⁴ For example, Massachusetts uses an enormously byzantine definition of the mental state for murder, without explaining the component parts:

Murder committed with deliberately premeditated malice aforethought, or with extreme atrocity or cruelty, or in the commission or attempted commission of a crime punishable with death or imprisonment for life, is murder in the first degree.

to the courts to define the very crimes themselves.⁵ Thus, jurors are required to somehow grasp these baffling notions having no prior understanding of the law, and render accurate verdicts. Adding perhaps the final touch to this, juries are “presumed” to follow their instructions, as well as understand the answers to whatever questions they put to judges during deliberations.⁶ Thus, judicial review rarely provides any correction for mistakes made.

Everything thus far discussed applies not only to ordinary cases, but to capital ones as well. Thus, since the United States was the only country in the Americas to execute someone in 2012,⁷ you would further have to explain the subtleties of death penalty law to your audience, including the roles played by factors in aggravation and mitigation.⁸ Naturally, since all of this is so daunting, this thought experiment leads to an irresistible conclusion: trust in the criminal jury system is problematic, as it is a system virtually rigged for failure.

But it is a system, and its flaws must be seen in that context.⁹ First, I will examine the developing literature on jury comprehension issues. Some speak in baleful terms here, talking of the “gloomy picture” presented and the distressingly low level of jury comprehension.¹⁰ However accurate those comments, jury issues are the result of various factors, many of which can be addressed successfully.

Murder which does not appear to be in the first degree is murder in the second degree. Petit treason shall be prosecuted and punished as murder. The degree of murder shall be found by the jury.

MASS. GEN. LAWS ch. 265, §1 (2011). Thus, we are left adrift, not only as to what is meant by murder itself, but as to what that four-word mental state means at all.

⁵ West Virginia is a pointed example, captioning one section “Voluntary manslaughter, penalty,” and providing no definition but only the penalty for whatever that crime is. W. VA. CODE § 612-2-4 (2011).

⁶ *Weeks v. United States*, 528 U.S. 225, 234 (2000) (“A jury is presumed to follow its instructions. Similarly, a jury is presumed to understand a judge’s answer to its question.”) (citations omitted).

⁷ See AMNESTY INT’L, DEATH SENTENCES AND EXECUTIONS 2012 (2013), available at http://www.amnesty.org.uk/uploads/documents/doc_23136.pdf.

⁸ In a series of cases beginning with *Furman v. Georgia*, 408 U.S. 238 (1972), the Court did not find the death penalty *per se* unconstitutional, but struck down statutes that mandated death, as they did not provide particularized consideration of individual defendants, and were incompatible with contemporary American values. *Woodson v. North Carolina*, 428 U.S. 280, 295 (1976).

⁹ Thus, Twain surely overstated things in calling the jury system “the most ingenious and infallible agency for defeating justice that human wisdom could contrive.” MARK TWAIN, ROUGHING IT 341(1872). For an excellent summary of how juries are being used worldwide, see Valerie P. Hans, *Jury Systems Around the World*, SCHOLARSHIP@CORNELL LAW: A DIGITAL REPOSITORY (Jan. 1, 2008), <http://scholarship.law.cornell.edu/cgi/viewcontent.cgi?article=1378&context=facpub>.

¹⁰ See Phoebe C. Ellsworth, *Are Twelve Heads Better Than One?* 52 LAW & CONTEMP. PROBS. 205, 219 (1989).

However, some see the criminal jury as an institution peculiarly suited to *deviating* from legal strictures to “make individualized moral judgments” that better serve the law than its strict application.¹¹ They are wrong, but nullification and other departures from the law must be discussed, representing a disturbing area of jury mistake-making.

Next, I will discuss several topics that confront the legal issues with which juries struggle. Juries are customarily instructed on the law only at the end of trials. Worse, those instructions often ignore their audience, frequently seeking legal accuracy over comprehensibility. And, when juries submit questions during deliberations, judges often simply re-read those puzzling instructions. Troubling as that is generally, it is completely unacceptable in capital cases. Worst yet perhaps, appellate review is terribly unaccommodating, again vaulting legal correctness over comprehensibility.

Finally, I will examine the codes themselves from which many of these problems spring. Some years ago, Paul Robinson posed a seminal question: Are criminal codes irrelevant?¹² Noting the enormous efforts devoted over the years to code reform, Robinson mused over whether any of this mattered, given the actual effect of codes on the systems they governed. They do matter, but they must be revised to eliminate the language and notions so ruinous to their real-world application, lest they get hopelessly lost in translation.

II. COMPREHENSION ISSUES

Jurors face overwhelming challenges. Chosen somewhat randomly from the general population, they are thrown into the fray with virtually no introduction to the tasks awaiting them. Then, with no real roadmap to follow, they sit on cases of varying complexity, to be asked ultimately to solve disputes with the only legal knowledge they have, the instructions given to them by the judge. And their decisions have obvious consequences, as in capital cases they are asked quite literally to choose between life and death in an area that has puzzled serious thinkers for millennia.

The entire process, from the trial’s beginning to possible appellate review bears a surreal cast. As Jerome Frank noted about the language of the instructions, “... everyone who stops to see and think knows that these words might as well be spoken in a foreign language, that indeed, for all the jury’s understanding of them they *are* spoken in a foreign language. Yet, every day, cases which have taken weeks to try are reversed by upper courts because a phrase or a sentence, meaningless to the jury, has been included in or omitted from the judge’s charge.”¹³

¹¹ Darryl Brown has written extensively in this vein. See, e.g., Darryl K. Brown, *Plain Meaning, Practical Reason, and Culpability: Toward a Theory of Jury Interpretation of Criminal Statutes*, 96 MICH. L. REV. 1199, 1209 (1998).

¹² Paul H. Robinson, *Are Criminal Codes Irrelevant?*, 68 S. CAL. L. REV. 159 (1994).

¹³ JEROME FRANK, *LAW & THE MODERN MIND* 195 (1963) (emphasis added).

Lost in Translation

Confusion is piled upon confusion, as jurors are asked to decipher inherently problematic terms, deploying poorly understood concepts in an unfamiliar setting.¹⁴ The literature on comprehension issues is rich, as for at least the last thirty years, social scientists have studied jury conduct. Though these issues subdivide rather subtly, virtually all studies note an unacceptable level of jury confusion on issues of substantive importance.¹⁵

For example, Professor Ellsworth conducted a study involving eighteen mock juries that watched a videotape of a homicide trial. After having been instructed, the jurors gave their initial verdict and were then assigned to juries to deliberate for one hour. Within one jury, the following colloquy took place (with Professor Ellsworth's comments):

Juror A: Second-degree stated that it doesn't have to be - he doesn't have to premeditate that far in advance." (scored as correct);

Juror B: If it's not premeditated, it can't be murder." (scored as incorrect; the jury accepts this);

Juror X: Now if I got up and I walked over there and you hit me, as I was coming over, then you hit me, then I pulled a knife out and stabbed you, that's manslaughter." (unclear);

Juror Y: No. That's self-defense." (unclear);

Juror X: Yeah, that's right. Self-defense would be manslaughter. (incorrect);

Juror Z: It's *involuntary* manslaughter." (incorrect).¹⁶

Obviously, these jurors had poor control over these concepts and, worse, the most forceful members held sway over the others.¹⁷ However, the very range of concepts considered showed just how challenging was the task they faced.

However, the cognitive challenges confronting jurors involve more than simply grasping the relevant legal notions or dealing with impenetrable legalese. They must perform a variety of tasks, some simultaneously, thus enhancing the potential for error. In an important early study, Nancy Pennington and Reid Hastie identified a series of tasks that jurors must perform:

¹⁴ See, e.g., the discussion of this by Judge Jon O. Newman in *Cabrera v. Jakobovitz*, 24 F.3d 372, 381 (2d Cir. 1994), referring to the inherently confusing nature of the "lawyerly cant."

¹⁵ See Ellsworth, *supra* note 10; Alan Reifman, Spencer M. Gusick & Phoebe C. Ellsworth, *Real Jurors' Understanding of the Law in Real Cases*, 16 LAW & HUM. BEHAV. 539 (1992); Shari Seidman Diamond, Beth Murphy & Mary R. Rose, *The 'Kettleful of Law' In Real Jury Deliberations: Successes, Failures, and Next Steps*, 106 NW. U. L. REV. 1537(2012) (discussing successes in an Arizona experiment in the context of disappointing results generally).

¹⁶ Ellsworth, *supra* note10, at 219.

¹⁷ *Id.* at 220.

- 1) The jury members must “encode” the information they get at trial. A competent jury must pay attention to the testimony and remember it.
- 2) The jury must define the legal categories. A competent jury should define these categories as they are presented in the judges' instructions.
- 3) The jury must select the admissible evidence and ignore evidence that is inadmissible.
- 4) The jury must construct the sequence of events.
- 5) The jury must evaluate the credibility of the witnesses.
- 6) The jury must evaluate the evidence in relation to the legal categories provided in the instructions. That is, certain elements of the story the jury constructs are particularly important in determining the appropriate verdict. The jury must identify these elements and understand how differences in the interpretation of the facts translate into differences in the appropriate verdict choice.
- 7) The jury must test its interpretation of the facts and the implied verdict choice against the standard of proof: preponderance of evidence, clear and convincing evidence, or beyond a reasonable doubt.
- 8) The jury must decide on the verdict.¹⁸

These tasks all involve dealing with law in one way or another, but through a variety of filters. Thus, whereas the jury will be instructed on ignoring inadmissible evidence, it must somehow sequence the events within certain legal constructs.

A case involving first-degree murder may, for example, require the jury to consider the time-frame of the killing, to decide whether the defendant had premeditated the killing and had an adequate period to reflect on that choice. Similarly, it might have to evaluate the same sequence of events to determine whether that killing was murder, or justified by the necessity of self-defense.

Though these tasks need not be performed sequentially, and though some may recede in importance depending on the case, the challenges are still obvious. Indeed, though many commentators seem to assume a rational decisionmaking model for juries, and may fault them for failing to behave with optimal rationality,¹⁹ Hastie has proposed a different model to

¹⁸ Nancy Pennington & Reid Hastie, *Juror Decision Making Models: The Generalization Gap*, 89 *PSYCHOL. BULL.* 246, 249-55 (1981) *paraphrased in* Ellsworth, *supra* note 10, at 207-08.

¹⁹ Certainly the murder trial of O.J. Simpson has received a great deal of attention of this sort. Indeed, Hastie and Pennington viewed this case from their storytelling perspective. Reid Hastie & Nancy Pennington, *The O.J. Simpson Stories: Behavioral Scientists' Reflections on the People of the State of California v. Orenthal James Simpson*, 67 *U. COLO. L. REV.* 957 (1996). There, the authors noted the challenges to a jury of a case

explain how juries work. Rejecting this model of optimal rationalism, he proposed that “the central process in jury decisionmaking is *story construction*—the creation of a narrative summary of the events under dispute.”²⁰

More than clever metaphor, the Story Model also captures how attorneys most persuasively appeal to juries, though the word “persuade” may be somewhat disingenuous. More accurately, the storytelling model provides a powerful vehicle for jury manipulation. Speculating that stories play a central role in the decisionmaking process, Hastie performed an experiment to determine whether the order of evidence presentation affected jury verdicts.²¹ He performed an experiment to determine the effect of variations in the order in which evidence was presented, as between “witness order” and chronological sequence.²²

His hypothesis was affirmed. Juries were more likely to convict when the prosecution evidence was presented in “story order” and the defense in “witness order.” Those results were reversed with a reversal of presentation orders.²³ Most significantly, these results bore no relationship to the relative strength of the cases. Though Hastie posits other non-rational factors in the decision making process, they just reinforce the central premise here: Jury verdicts can be skewed by a variety of factors bearing little or no relationship to the abstract strength of cases or perceived legal truths.²⁴ Indeed, the power of effective storytelling is dramatically demonstrated by a case, conceived by many to be a “dead-bang winner” for the state, the case of the famous subway vigilante, Bernhard Goetz.²⁵

III. NULLIFICATION AND OTHER DISTORTIONS OF THE LAW

On December 22, 1984, Bernhard Goetz had the sort of urban encounter many fear on a daily basis. Approached by four teenagers in a crowded subway car, he refused to submit to their demand that he give

marked by largely circumstantial evidence, serious problems of the meaning of reasonable doubt, both in the context of a trial in which the jury was sequestered for over a year. *Id.* at 958. By their view, though, the story of the racist police detectives and bumbling criminalists swayed the jury to acquit.

²⁰ Reid Hastie, *Emotions in Jurors' Decisions*, 66 *BROOK. L. REV.* 991, 995 (2001). This has commonly become called the story model, and Hastie uses that term in this article and elsewhere.

²¹ As he said, the objective was to determine whether the stories played this central role, as opposed to being “merely constructed ‘on the side,’ while the actual decision process relied on other cognitive processes and conclusion.” *Id.* at 997.

²² *Id.*

²³ *Id.* at 998.

²⁴ Hastie’s article is, after all, called *Emotion*, and he has spent a lengthy career writing about decisional models and juries.

²⁵ Goetz was dubbed the subway vigilante by the New York Post, after having been described that way by then mayor Ed Koch. GEORGE P. FLETCHER, *A CRIME OF SELF-DEFENSE: BERNHARD GOETZ AND THE LAW ON TRIAL* 2(1988)[hereinafter FLETCHER].

them five dollars.²⁶ Instead, he removed his .38 revolver from his waistband, and emptied the five-shot Smith & Wesson into them. Seeing that he had missed one, he pursued him to where he sat cowering on a seat, saying “You seem to be all right; here’s another.”²⁷ With that shot, he severed the spinal cord of Darrell Cabey. After the subway screeched to a halt, he had a brief conversation with the driver and fled into the night.

From there, he drove to New Hampshire, where he eventually made a videotaped confession, and was turned over to the New York police on January 3. A grand jury failed to indict him on the assault charges several weeks later, as none of the victims and few witnesses appeared. However, as his celebrity grew, pressure mounted to resubmit the assault charges, and a grand jury finally indicted him on ten charges of assault on March 27, 1985.

Two years later, on March 23, 1987, the trial began. Most onlookers, myself included, expected a defense solely tethered to self-defense. In his rambling confessions, a clearly disturbed Bernhard Goetz squarely admitted his desire to kill the youths:

I wanted to kill those guys. I wanted to maim those guys. I wanted to make those them suffer in every way I could -- and you can't understand this because it's a realm of reality that you're not familiar with. If I had more bullets I would have shot them all again and again. My problem was I ran out of bullets.²⁸

Since the mental state for attempted murder is invariably the intent to kill, and since in both action and word Goetz intended to kill those four, the state’s case appeared invincible. But as Fletcher pointed out, it was not truth alone that was on trial, but rather it was about “the kinds of people who confronted each other that Saturday afternoon.”²⁹ Goetz was a troubled, beleaguered middle-aged man who had been previously mugged and seriously injured. Because of bureaucratic bungling, he had been denied a gun license, yet frequently travelled the city in anger and fear, his gun at his waistband. The youths all had extensive criminal records and their “boisterous” behavior that day had already prompted many riders to leave that part of the subway car.³⁰ The stage was clearly set for defense counsel to capitalize on that imagery, and they readily obliged.

Whereas the Simpson defense was constructed around the imagery of corrupt LA police and bumbling criminalists, Goetz’s played on those of urban fear, its story unmistakable and familiar. Central to the story was

²⁶ *Id.* at 1.

²⁷ *Id.*

²⁸ *The Trial of Bernhard Goetz: Goetz’s Videotaped Confession*, UNIVERSITY OF MISSOURI-KANSAS CITY SCHOOL OF LAW, <http://law2.umkc.edu/faculty/projects/ftrials/goetz/goetzconfession.html> (last visited June 4, 2013).

²⁹ FLETCHER, *supra* note 25, at 102.

³⁰ *Id.* at 3.

the suffocating image of someone trapped in a situation with no good resolution.³¹ His attorney, Barry Slotnick, directed a defense that was a meticulously scripted play, right down to the casting of four street-clad Guardian Angels as the four teenagers who confronted Goetz.

Surrounding the “Goetz” character in the courtroom, presumably to demonstrate the paths of the bullets, they communicated a sense of fear “more powerful than rational argument.”³² There, and elsewhere in the trial, Slotnick resorted to dramatic storytelling, defending his client against technically overwhelming charges with surpassing powerful emotion. Storytelling at its best, this approach effectively pushed law to the side.

Initially focusing on technical legal requirements, the jury eventually veered off course completely. Though self-defense seemed to be the only basis for exoneration, the jury never reached that issue. Yielding to the persistent appeals to their moral sensibilities, they started to blur self-defense and mental state. One juror questioned the state’s proof of his intent to kill, saying “[w]e needed a motive for murder. The only motive that Waples presented to us was this revenge. And we didn’t buy it.”³³ Thus, seeing Goetz as reacting to *their* aggression, they reconceptualized “intent” as somehow requiring an evil motive. Seeing none, they acquitted him on all charges of attempted murder and assault, including that on Cabey, whom he shot at twice. In their eyes, he was not morally blameworthy and thus not guilty.

But consider what the jury would have done had Goetz *actually killed someone*. Wouldn’t a conviction have been inevitable, assuming the same conduct, including his various confessions? Could the jury plausibly have viewed him as morally correct, as it did in the actual case? Perhaps the verdict simply reflects the fact that common sense prevailed, as no one was killed, and the label “victims” seemed a complete misnomer for those teenagers. Or so Fletcher thought.

Defending this result, Fletcher saw it as “perfecting the law.” For him, “careful historical reflection underscores the power of the jury not to defeat the law, but to perfect the law, to realize the law’s inherent values.”³⁴ In fact, resisting the term nullification, Fletcher justified the jury’s prerogative to ignore law’s technical demands, presumably to better serve its fundamental calling. Thus, for Fletcher and others “carefully constructed legal edifices crumble at the touch of the jury’s common sense.”³⁵

But that is strange. Had the jury considered self-defense head on, it would have dealt with whether Goetz had acted reasonably under the circumstances. Though the use of that malleable standard, it would have con-

³¹ As Fletcher said, “it is difficult to find a New Yorker, black or white, inclined to believe that Goetz could have avoided a violent confrontation simply by ignoring Canty’s request (demand?) for money or deflecting his request with an offhand remark....” *Id.* at 101.

³² *Id.* at 130.

³³ *Id.* at 186 (statements of juror D. Wirth Jackson).

³⁴ *Id.* at 154.

³⁵ *Id.* at 188.

sidered the self-same facts that led it instead to usher in self-defense through the back door. But in behaving that way, hadn't it rejected its duty to apply the law? Wasn't that its only task?

A. JURY NULLIFICATION AS JURY ERROR

Statutes are inherently general, as they set standards that apply to a broad swath of the population, despite substantial personal differences and backgrounds within that group. Thus, it's unremarkable that bedrock notions such as reasonableness and mental states apply unevenly across different demographics. Presumably, that's what Fletcher meant as he described the Goetz jury as "perfecting the law."

The notion of crime commission may assume a proactive actor, one who initiates a problem unprovoked and unjustified. That's an imperfect fit for Goetz. A sad little man, he was just trying to make it through his day, getting from here to there unbothered. Thus, the inevitability of further harassment by those kids probably haunted the jury, as the looming question was always one of "what would you have done" in a similar confrontation.³⁶ It is easy to see how we can, then, slip into thinking that an imperfect law calls for fine-tuning, or some other form of correction by caring juries.

A criminal conviction obviously carries serious consequences, and many think that, standing between the state and the individual, the jury has and should exercise the power to make these individualized moral judgments as they see fit. Reflecting a similar dissatisfaction with the limited nature of criminal conduct rules, Brown noted that they "inevitably lack the nuance to control fully the particularized moral judgment of a defendant's conduct in his specific context."³⁷ This makes an appealing case for a jury's exercising this function in assessing criminal liability.

Supporting his argument, Brown presented an appealing case that straddled the line between interpretation and nullification. With permission of a Wisconsin state court, PBS videotaped the deliberations of a jury in a simple case involving possession of a firearm by a convicted felon. The facts were not in dispute, but the law-facts fit was.

A man of "substantially sub-average intelligence," Leroy Reed, clearly possessed a gun in violation of state law.³⁸ He read at a second-grade level, and had aspirations to become a private detective. He explained that to a police officer whom he met in the courthouse, and the officer asked him to

³⁶ Indeed, this is a constant refrain of Fletcher's, as he considered the unfortunate choices facing Goetz. Would simply showing the gun have deterred the kids? How about if Goetz had simply ignored them? These and similar notions apparently led Fletcher to conclude that the jury's task was not to solve a factual dispute, but to invoke higher values. *See, e.g., id.* at 101.

³⁷ Brown, *supra* note 11, at 1209. Elaborating in a footnote, Brown referred to the extensive literature dealing with the need for a conciseness in criminal statutes that necessarily limits the range of factors applicable to guilt.

³⁸ *Id.* at 1239.

produce the gun by bringing it to the police department. He did so and was arrested.³⁹

Though the judge did not instruct the jury on nullification, the “story” of the case was clear: Reed was simply a slow man, presenting little threat, ensnared by overzealous police and prosecutors. To convict him, would compound that injustice. As a result, the jurors quickly considered their appropriate role. One, commenting that it is appropriate for juries to judge the law itself, openly questioned whether he should just apply the elements in a “cut and dried” fashion.⁴⁰ Echoing similar sentiments, another, a school psychologist, questioned whether Reed “knew” he possessed a gun with sufficient depth to be blameworthy, asking, “at what *level* did he know” that.⁴¹ Arguably, the jury was within bounds in simply seeking to apply the law competently, but it’s less clear whether that was happening, or whether, seeing him as a harmless and sympathetic character, it simply resisted attaching blame. It acquitted him within two and a half hours of deliberation.

However, it *is* clear that the jury felt it was doing justice in avoiding what it saw as an unacceptably strict application of the law. Its deliberations were rife with such comments, one juror explicitly questioning “[a]re we obligated ... to follow the letter of the law and find him guilty, or are we obligated as a jury to use our special level of conscience.”⁴² There, that struck the note most frequently sounded in these discussions, the sense that the jury represents the conscience of the community and is thus asked, if not obligated, to reflect its normative beliefs. But should that be?

Reprising that conversation with those foreign lawyers, imagine their reaction to the phenomenon of jury nullification. Surely nullification is inevitable in a system such as ours, but they might be stunned at the romantic folklore surrounding this, with its frequent invocations of John Peter Zenger and the like.⁴³ But nullification is dangerous on several counts.

First, nullification discussion assumes juries that certainly get it on the law. Having been charged, having understood those instructions and having successfully held the law up against the facts, they somehow choose to resist the law. That resistance creates a tension between the populist mythology of nullification and calmer notions of the jury function. Laws are enacted, often with great difficulty, through the process of representative democracy. For example, the notion of willful blindness has sharply divid-

³⁹ *Id.* at 1240.

⁴⁰ *Id.* at 1242.

⁴¹ *Id.* at 1245(emphasis added).

⁴² *Id.* at 1242.

⁴³ See JAMES ALEXANDER, A BRIEF NARRATIVE OF THE CASE AND TRIAL OF JOHN PETER ZENGER (1736). Few ancient cases are dealt with more frequently and with greater reverence than this American chestnut.

ed many legal commentators.⁴⁴ When pressed to decide whether and how to legislate on this, our representatives have a great deal to consider. But they *are* representatives and the process may be appropriately arduous. Not so with jury nullification.

The *Reed* jury acquitted after scarcely two hours.⁴⁵ Presumably, some of that time was spent on preliminary and formal matters, such as electing a foreperson. Then, in an unstructured process, perhaps yielding to its more forceful members, the body decided to acquit, perhaps happy to do so given the inoffensive defendant and lure of an early end to jury service. Thus, this strange form of law creation takes place *ad hoc* by an unaccountable, perhaps unrepresentative, group that may not even reflect community sentiment (if such exists). But the problems run deeper.

Nullification's history reflects our national repugnance at tyrannical regimes. But the phenomenon also reflects an acknowledgement of the limitations of bright line tests for blameworthiness. Perhaps *Reed* was slow and an inappropriate candidate for that prosecution. Yet in acquitting, the jury implicitly decided it knew more, or knew better, than those in the system who moved the case along. Thus, the first erroneous assumption supporting this power is that jurors somehow know enough to make reasoned decisions to nullify.

As Andrew Leipold noted, juries may act on unsubstantiated, superficial impressions, bereft of evidentiary support. Thus, as he pointed out, a jury might well acquit a clean-cut student of drug possession, without knowing anything more about him precisely *because* of the limits placed on criminal trials.⁴⁶ The decision, presumably well tailored to assess individual blameworthiness might, rather ironically, miss the mark entirely. But there is a more troubling problem here, one that more likely might turn the public conscience justification on its head.

Presumably, the impetus for nullification comes from the jury's desire to do justice by showing mercy or lenience toward that individual defendant. But that can backfire horribly.⁴⁷ Though we assume law lags behind social awareness on many issues, it can be just the opposite. A jury might nullify in a case of domestic violence, thinking the problem was a private one. Indeed, that might draw from a libertarian sentiment favoring privacy

⁴⁴ Willful blindness has been referred to as a "form" of knowledge, though more often as a kind of ostrich-like ignorance. See, for example, the majority and dissenting opinions in *United States v. Jewell*, 532 F.2d 697 (9th Cir. 1976).

⁴⁵ For fascinating synopses of the proceedings, see *But Did He Know It Was a Gun?: A Paper for the International Pragmatics Association Meeting*, BRADLEY UNIV. (July 5, 1996), <http://bradley.bradley.edu/~ell/leroy.html>. Those involved in the production for PBS presented a paper delivered at the International Pragmatics Association Meeting in 1966. That is a Belgian group devoted to the study of language use and communication. See INT'L PRAGMATICS ASS'N, <http://ipra.ua.ac.be/> (last visited Aug. 12, 2013).

⁴⁶ Andrew D. Leipold, *Rethinking Jury Nullification*, 82 VA. L. REV. 253, 303 (1996).

⁴⁷ *Id.* at 304-05.

over government intervention. However, that result might reflect an ignorance in the area, one not shared by legislators.

The jury might be unaware of the social costs and public harms resulting from domestic violence.⁴⁸ Perhaps drawing from limited experience or pop notions, it might see prosecution as invasive in an area best left untouched by public officials. Acting out of compassion, it might totally ignore the costs to our healthcare system as well as to all involved in this sordid phenomenon. Not acting out of crassness or indifference, it might nevertheless reflect a regressive mentality inconsistent with current knowledge in the area.

Finally, as Leipold also points out, nullification may be incorrectly bottomed on the notion of law's inflexibility, its failure to account for the outlier case.⁴⁹ Assume the Goetz jury had first discussed self-defense, thinking that was the only real issue there.⁵⁰ Assume further that it had to follow the rigorous course of asking who the aggressor was and similar questions. It would then have considered whether he had reasonably used lethal force, and would have been forced to consider his alternatives. If it found that he acted precipitously, it would probably have convicted.⁵¹ That wouldn't have been based on law's inflexibility, but rather on the fact that Goetz acted wrongly in the very face of an extremely flexible defense, one that forces the fact finder to view the actor's conduct contextually.

It's fanciful to think that all nullification can be eliminated. It is a power and thus, even without the benefit of instructions on it, juries nullify. However, if jury deliberations are better structured, if the judge plays a more active, instructive role in its conduct, perhaps the incidence of nullification can be substantially reduced. And, surely jury errors result from an often stunning judicial passivity and unwillingness to better structure jury conduct.

⁴⁸ The legislative process frequently involves a fact-finding process aided by experts during legislative hearings. For example, Tennessee is currently holding hearings on these costs. See Lindsay Burkholder, *Tennessee Panel Looks at Cost of Domestic Violence*, TIMES FREE PRESS (May 1, 2013), <http://www.timesfreepress.com/news/2013/may/01/tennessee-panel-looks-at-cost-of-domestic-violence/>. In those hearings, legislators can learn of the obvious and indirect costs resulting from this violence. See *Community Costs of Domestic Violence*, STOP VIOLENCE AGAINST WOMEN: A PROJECT OF THE ADVOCATES FOR HUMAN RIGHTS, http://www.stopvaw.org/community_costs_of_domestic_violence (last updated July 19, 2011)(detailing the frightening costs, both direct and indirect, of domestic violence).

⁴⁹ Leipold, *supra* note 46, at 308.

⁵⁰ Again, recognize the striking absence of any fixed format or process for jury deliberations. Accordingly, this is a very sensible assumption.

⁵¹ Indeed, had it rejected self-defense, it would have been hard-pressed to usher it in through the back door, as it did.

IV. EDUCATING JURIES

Resistance to overbearing law by juries reveals something important: They take their work very seriously and recognize the consequences of their actions. Indeed, virtually all studies of jury behavior even show that judges overwhelmingly agree with jury verdicts, further demonstrating their seriousness of purpose.⁵² But those numbers can be looked at differently; assuming that agreement means that the juries were correct, that still means that they are *wrong* perhaps 25% of the time. That must be addressed.

The last half-century has seen the emergence of pattern jury instructions throughout the country. With that, there is greater consistency within jurisdictions and greater thought given to the process of educating juries. But there remains a disquieting gap between legal accuracy and effective communication in much of this. Thus, for example, Alabama has had pattern instructions for forty years now, but only recently acknowledged that “though pattern instructions are true to the law ... even ‘scholarly,’ the instructions are complicated definitions of legal principles.”⁵³ California recently re-wrote its civil and criminal instructions, but earlier wrote remarkably that the thing “an instruction must do above all else is correctly state the law. This is true regardless of who is capable of understanding it.”⁵⁴ That concession to understanding is painfully evident.

A well-known Supreme Court case may best illustrate this overwhelming challenge facing a lay jury. In *Martin v. Ohio*, the Court confronted a case involving the intersection of several difficult issues. There, a wife who concededly killed her husband claimed self-defense.⁵⁵ She was charged with aggravated murder, which, under Ohio law consisted of “purposely, and with prior calculation and design, causing the death of another.”⁵⁶ Obviously, the state had the burden of proving that beyond a reasonable doubt.

But she also bore a burden, not only of proof, but also of going forward with evidence in this affirmative defense setting. To prove self-defense, she had to prove its elements by a fair preponderance of the evidence. But here’s the rub: under Ohio law, self-defense requires proof of three items. First, the defendant must show that she was fault-free in creating the fray, conventionally, that she was not the aggressor. Second, she had to show she honestly believed there was either a lethal threat or serious injury threatened, from which only lethal force could protect her. Finally,

⁵² See Diamond et al., *supra* note 15, at 1543 (citing numerous studies showing that judges agree with juries in roughly 75% of criminal cases).

⁵³ Arthur J. Hanes, JR., Bert S. Nettles & Leilah Watson, *The ‘Plain English’ Project of the Alabama Pattern Jury Instructions Committee-Civil*, 68 THE ALA. LAWYER 368, 369 (2007), available at http://www.alabar.org/publications/articles/Sept07/plain_english.pdf (last visited January 20, 2014).

⁵⁴ Robert L. Winslow, *The Instruction Ritual*, 13 HASTINGS L.J. 456 (1962) (quoting 1 BAJI, CAL. JURY INSTRUCTIONS — CRIM. BOOK OF APPROVED JURY INSTRUCTIONS 44 (1950)).

⁵⁵ *Martin v. Ohio*, 480 U.S. 228 (1987).

⁵⁶ *Id.* at 230.

she had to show the necessity of using that force by showing the unavailability of retreat or desistance.⁵⁷

So consider the challenge to the jury thus far. On the substance, it had to juggle the notion of the defense's burden of proof and production on self-defense, using the preponderance standard. Then, it had to align that against the state's burden of proof of the elements. To do this, it had to grasp any potential *overlap* between those elements and recognize the overarching burden of the state to prove guilt. Without careful guidance, that's well-nigh impossible. Understand, so long as the defendant's proof of her defense raises a reasonable doubt on one of these overlapping issues, she must be acquitted. She is not required to go further and meet that preponderance standard.⁵⁸ Since this was a Due Process challenge, an instruction that arguably shifted the burden from the state violated her rights.

The Court recognized this, acknowledging, "[t]he instructions in this case could be clearer in this respect."⁵⁹ However, it went on to conclude (somehow) that they were sufficient, "read as a whole."⁶⁰ That is one of those puzzlingly odd empirical observations sometimes made by the Court, one that ignores the real-life context in which events transpire. For example, with no set order in which to consider matters, the jury might have first considered self-defense, as she clearly killed him. Had she not met her required standard of proof, it might have then *dismissed* that constellation of issues, and moved on to the state's case. At that point, only the rare, remarkably astute jury member would consider any overlap between the two and point out the problem.⁶¹ Justice Powell made that point, as he lamented this weakened presumption of innocence that resulted from this decision.⁶²

Convicted of aggravated murder, she could have faced the death penalty, following an instruction that could create deadly confusion, one wors-

⁵⁷ *Id.*

⁵⁸ Despite our hope for clarity in writing these pieces, I almost hope this is puzzling. Imagine, then, the plight of the jury in such a case!

⁵⁹ *Martin*, 480 U.S. at 234.

⁶⁰ *Id.*

⁶¹ Indeed, as Joshua Dressler asks in his casebook: "if *you* were a juror, would you understand who has the burden of persuasion—and quantum of proof—in regard to: (a) whether M killed V; (b) whether M killed V purposely; (c) whether M killed V with prior calculation and design; (d) whether M was at fault in the difficulty; (e) whether M should have retreated before killing V; and (f) whether M honestly believed that she was in imminent danger of death or great bodily harm?" JOSHUA DRESSLER, *CASES AND MATERIALS ON CRIMINAL LAW 477* (6th ed. 2012).

⁶² *Martin*, 480 U.S. at 237-38 ("The reason for treating a defense that negates an element of the crime differently from other affirmative defenses is plain. If the jury is told that the prosecution has the burden of proving all the elements of a crime, but then also is instructed that the defendant has the burden of *dis* proving one of those same elements, there is a danger that the jurors will resolve the inconsistency in a way that lessens the presumption of innocence.").

ened by judicial reluctance to provide any clarity.⁶³ *Weeks v. United States* presents this somber possibility.

At the death penalty phase of his trial, the prosecution proved two aggravating circumstances and the defense called ten witnesses to prove mitigation.⁶⁴ During deliberations, the jury asked the judge whether it was their duty to impose death if it found that the state successfully proved aggravation.⁶⁵ The instruction read:

If you find from the evidence that the Commonwealth has proved beyond a reasonable doubt either of the two alternatives, and as to that alternative you are unanimous, then you may fix the punishment of the defendant at death or if you believe from all the evidence that the death penalty is not justified, then you shall fix the punishment of the defendant at life imprisonment or imprisonment for live [*sic*] and a fine of a specific amount, but not more than \$100,000.00.⁶⁶

The problem was clear: perhaps unaccustomed to dealing with permissive language (“may”), and certainly unfamiliar with death penalty law, the jury was understandably confused about its duties. In a written response, the judge referred them back to the instruction reproduced above, stating: “I don’t believe I can answer the question any clearer than the instruction, so what I have done is referred them” to its second paragraph.⁶⁷

The jury sentenced Weeks to death two hours after this reply. Following the conviction, he appealed within the Virginia system, and petitioned for *habeas* in federal court. Losing at every stage, he sought relief in the Supreme Court. That Court’s conclusion was clear from its issue statement: it asked “whether the Constitution is violated when a trial judge directs a capital jury’s attention to a specific paragraph of a constitutionally sufficient instruction in response to a question regarding the proper consideration of mitigating circumstances.”⁶⁸

The Court occupied an odd position. It couldn’t know whether the jury eventually got the jury instructions, or simply gave up and arrived at its verdict. Moreover, there was no substantive deficiency in that instruction, at least in its accuracy; but, however understandable the Court’s focus on accuracy of content, it still lost sight of its mission, to determine whether a jury erroneously condemned a man to death. Thus, following the “presumption” that juries follow their instructions and the additional presumption that it understands the answers to the questions it poses, the Court narrowly affirmed the conviction, noting only a “slight possibility” that the

⁶³ To this day, Ohio is one of thirty states that still retain the death penalty. Death Penalty Information Center, found at *States With and Without the Death Penalty*, DEATH PENALTY INFO. CENTER (2013), <http://www.deathpenaltyinfo.org/states-and-without-death-penalty>.

⁶⁴ *Weeks v. Angelone*, 528 U.S. 225 (2000).

⁶⁵ *Id.* at 228.

⁶⁶ *Id.* at 229.

⁶⁷ *Id.*

⁶⁸ *Id.* at 227.

jury was confused.⁶⁹

Dissenting, Justice Stevens focused on communication, troubled by the potential for “erroneous *interpretation*” of facially correct instructions.⁷⁰ To those Justices,⁷¹ the record established a “virtual certainty” that the jury did not apply the law correctly.⁷² Accustomed as we are to seeing ideological splits within the Court, this is more. Justice Stevens pushed the analysis beyond presumptions, to try to determine when possible jury confusion warrants reversal.⁷³ The Court’s split perfectly mimics this dichotomy of accuracy and clarity, in a trying area for the judiciary. Weeks was executed two months after this decision.⁷⁴

Steve Garvey decided to put these widely differing assumptions to the test.⁷⁵ His team placed ads in two Virginia newspapers for participants in mock sessions on capital sentencing. A total of 154 members of those communities participated, largely from the local college populations. They were asked several questions, largely focusing on the mandatory effect to be given the factors in aggravation.⁷⁶ Thus, in several settings, they were asked whether they felt the law required the imposition of death.

In the group that received the instruction from *Weeks*, a substantial

⁶⁹ *Id.* at 235. I highlight the word presumption, because of the unusual usage. See Judith L. Ritter, *Your Lips are Moving ... But the Words Aren't Clear: Dissecting the Presumption that Jurors Understand Instructions*, 69 MO. L. REV. 163, 166 (2004)(discussing the origin of this odd usage and its history in Supreme Court cases, as it does not function as a presumption, but as an assumption in the law).

⁷⁰ *Weeks*, 528 U.S. at 238 (emphasis added).

⁷¹ Stevens was joined by Justices Ginsburg, Breyer & Souter.

⁷² *Weeks*, 528 U.S. at 238.

⁷³ The dissent noted four aspects of the record which, taken cumulatively, warranted reversal: The text of the instructions, the judge’s responses to questions, the verdict forms and the court reporter’s transcript of the polling of the jury (apparently, many members were in tears at that point). *Id.* at 238.

⁷⁴ *Trooper’s Killer Executed in Virginia*, ASSOCIATED PRESS (Mar. 17, 2000), available at <http://www.washingtonpost.com/wpdyn/content/article/2010/03/16/AR2010031601702.html>.

⁷⁵ Stephen P. Garvey, Sheri Lynn Johnson & Paul Marcus, *Correcting Deadly Confusion: Responding to Jury Inquiries in Capital Cases*, 85 CORNELL L. REV. 627 (2000) [hereinafter Garvey]. As they put it, if jurors ask a judge for clarification, “chances are good they didn’t understand the instruction. Why else would they have asked the question.” *Id.* at 627.

⁷⁶ These were the two factors from the *Weeks* case: The first was the “probability that he would commit criminal acts of violence that would constitute a continuing threat to society” in the future. *Weeks*, 528 U.S. at 228. The second required a finding that his criminal conduct in committing the offense was “outrageously or wantonly vile, horrible or inhuman, in that it involved depravity of mind or aggravated battery to the victim beyond the minimum necessary to accomplish the act of murder.” *Id.* The Court did not quote the full text of the pattern instruction, but it is found at 4 RONALD J. BACIGAL & JOSEPH S. TATE, VA. PRAC. SERIES § 76.5 (2012) (capital murder).

minority felt a death sentence was compelled.⁷⁷ Anticipating that some might question the behavior of mock jurors, Garvey compared these results to those accumulated by the Capital Jury Project (CJP). CJP interviewed 650 jurors from seven states, and its results were strikingly similar, with only a minor deviation on one factor.⁷⁸

Next, Garvey replicated the situation from *Weeks*, simply redirecting the jury to the pattern instruction. Perhaps like the actual *Weeks* jury, the numbers of those who felt bound to apply the death penalty *actually increased* when simply asked to re-read the original instructions.⁷⁹ Naturally, that can't account for the actual dynamic of the jury behavior in *Weeks*, but it is still disturbing, since *Weeks* had killed a state police officer, and that jury focused on the aggravating factor of heinousness.

However, despite not knowing what happened in that actual jury, these numbers strongly substantiate Stevens' view that the very asking of the question revealed confusion, a confusion that undoubtedly remained after they simply re-read what was already problematic. Recall, the judge's reply to the question was simply "see second paragraph of Instruction #2
...."⁸⁰

Essentially, the defense requested the addition of two words to the pattern instruction, the words "even if." Thus, it asked the judge to instruct the jury that "[e]ven if you find that the State has proved one or both of the aggravating factors beyond a reasonable doubt, you may give effect to the evidence in mitigation by sentencing the defendant to life in prison." That instruction would not have skewed matters either way, and fully comported with the shared understanding of death penalty law.

In Garvey's control situation, he then gave that requested instruction to the mock jurors who were aware of the question sent from the actual jury. The clarification dispelled the confusion among forty percent of those who initially felt the death penalty was mandated.⁸¹ Finally, Garvey noted from his data that the jurors "who understood the rule were in fact more likely to vote for life compared with jurors who misunderstood the instruction."⁸² Since *Weeks* was probably a close case, the result could have turned on the addition of that two-word clarification.

But *post mortems* alone are pointless. Had this story been recounted

⁷⁷ Forty-one percent felt death was mandatory for the second factor, and thirty-eight percent for the first. Garvey, *supra* note 75.

⁷⁸ It had the same forty-one percent on factor two, and thirty-two percent on the first. Garvey, *supra* note 75, at 637.

⁷⁹ *Id.* at 638-39.

⁸⁰ *See id.* at 655 (app. VI).

⁸¹ Again, the results were slightly different as between the two aggravating circumstances. *See* Garvey, *supra* note 75. On heinousness, the numbers dropped from forty-nine percent to twenty-nine percent. *Id.* On future dangerousness, they went from forty-five percent to twenty-four percent. *Id.* Those numbers are still disturbing, as the requested charge was utterly clear. However, its elimination of substantial confusion is still enormously significant.

⁸² Garvey, *supra* note 75, at 641.

to our attentive foreign lawyers, someone would probably have asked *when* this innocuous but crucial correction was made to the law. It never was. Darryl Brown has written about the “decision effect” of jury instructions.⁸³ Equally accurate instructions can produce markedly different outcomes. Garvey’s findings corroborated that. Yet, because of a “‘failure in the market’ for jury instructions,” the system remained with sufficient, but greatly suboptimal instructions.⁸⁴

Perhaps somewhat unaware of linguistic and normative dimensions to this problem, the authors to the Virginia instructions never made that necessary, hoped-for change. Accordingly, the current version of the instruction exists unchanged from that given in *Weeks*, as “adapted from the language and instruction *approved* and given in *Weeks*.”⁸⁵ That instruction was “approved” only in the sense that a majority agreed that it met minimal constitutional standards.

Simply contrasting the notions of clarity and accuracy advances few causes, however. Rather, the focus must tighten to determine systemically how to eliminate instructions that, though accurate, are linguistically confusing and normatively undesirable.

V. EDUCATING JUDGES

The emergence of pattern instruction is generally welcomed. After all, fastidiously written by scholarly groups, they create accuracy and consistency within the judiciary. But they often do that at the expense of clarity, thus undermining their very purpose. Written for a judiciary often wary of error, they frequently do little of real value.

This can be seen in how some instructions have addressed two topics of enormous concern here: burden of proof and the role of factors in death penalty deliberations. In everything I’ve discussed, the jury must somehow grasp and apply the reasonable doubt standard. Familiar in pop culture, it can prove vexing to apply. Thus, until quite recently, the Tennessee pattern instruction read:

Reasonable doubt is that doubt engendered by an investigation all the proof in the case and an inability, after such investigation, to let the mind rest easily as to the certainty of guilt. Reasonable doubt does not mean a captious, possible or imaginary doubt. Absolute certainty of guilt is not demanded by the law to convict of any criminal charge, but moral certainty

⁸³ Darryl K. Brown, *Regulating the Decision Effects of Legally Sufficient Jury Instructions*, 73 S. CAL. L. REV. 1105, 1109 (2000). As he said, “when two instructions are both legally sufficient, advocates have no legal grounds on which to argue. Decision effect is unregulated by law and is, therefore, left to trial judges’ discretion.” *Id.* (footnotes omitted).

⁸⁴ *Id.* at 1118. Brown explains that judges choose among the available instructions (analogized to products), based on the best combination of accuracy and cost (reversal), often choosing an inferior one. *Id.*

⁸⁵ BACIGAL & TATE, *supra* note 76, at § 76:5 (emphasis added).

is required, and this certainty is required as to every proposition of proof requisite to constitute the offense.⁸⁶

Apparently accurate, it nevertheless fails to communicate with the modern juror. Indeed, the failing of many current instructions, this reflects a legal inertia in which change proceeds glacially. However, rather recently, lawyers trained in linguistics have started to focus on the flaws of pattern instructions, often aiding in the creation of so-called “plain English” instructions.⁸⁷

But that is just the beginning; parsing this Tennessee instruction, Bethany Dumas found numerous and varied flaws.⁸⁸ Structurally, she pointed out that those three sentences contain eleven clauses, “embedded at a C level.”⁸⁹ In addition, the use of arcane diction such as “captious” coupled with negative concepts (“inability”) and passive constructions (“demanded by the law”) further strain understanding. Worse, add to that the manner of delivery of the instructions (perhaps merely read to the jury), and we have an almost foolproof recipe for misunderstanding. Indeed, recognizing that educating juries is indeed an educational exercise, she noted the imbalance between what she called “domain experts.”⁹⁰

That is, the trial judge is the expert on the law, the jury on the facts. Despite that, there is an imbalance in the respect afforded the two groups. Jurors play an exceedingly passive role, one hardly advancing the educational function pursued. In all, in both structure and substance, this seems, as Phoebe Ellsworth noted, like a system “set up to promote misunderstanding.”⁹¹

And these jury tasks do not exist in isolation. A jury must *apply* this burden to the material elements of the offense. Surely the *Weeks* instruction strained that jury beyond comprehension. However, its obscurity on substance is easily matched by the simply dreadful craftsmanship of some instructions, thus further rigging the system for failure. Illinois only recently abolished the death penalty and had a pattern instruction on the role of factors, noting: “If you do not unanimously find from your consideration of all the evidence that there are no mitigating factors sufficient to preclude imposition of a death sentence, then you should sign the verdict requiring

⁸⁶ TENN. JUDICIAL CONFERENCE, PRAC. SERIES: TENN. PATTERN JURY INSTRUCTIONS – CRIM. § 2.03 (4th ed. 1995). Unfortunately its successor scarcely does better, essentially only removing the word “captious.” TENN. JUDICIAL CONFERENCE, TENN. PRAC. SERIES: TENN. PATTERN JURY INSTRUCTIONS – CRIM. § 2.03 (15th ed. 2012-13).

⁸⁷ Indeed, though Diamond hailed California as the only jurisdiction to attempt plain English rewrites, others also have. Diamond et al., *supra* note 15 at 1545. See, e.g., the Alabama project referred to in *supra* note 15.

⁸⁸ Bethany K. Dumas, *Jury Trials: Lay Jurors, Pattern Jury Instructions, and Comprehension Issues*, 67 TENN. L. REV. 701, 726-27 (2000).

⁸⁹ *Id.* at 726. By “C” level, she meant that you had to drill through two outline levels to reach matters such as “the mind rest easily as to the certainty of guilt.” *Id.*

⁹⁰ *Id.* at 727.

⁹¹ See Ellsworth, *supra* note 10, at 224.

the court to impose a sentence other than death.”⁹²

Devoid of arcane diction or obscure content, this brief instruction with its four negatives still completely fails to communicate to anyone, much less to a lay jury, on this critical topic. However, as I have said, both social scientists and psycholinguists are advancing this plain English movement, creating real promise for improving jury comprehension.

A. *EDUCATING JURIES BETTER*

Jury performance can be improved, despite the inertia in the system. As discussed, judges favor accuracy at all costs, and little change is likely from either the trial bench or its appellate counterpart. However, change may advance slowly, spurred by participants in the plain English movement.

Naturally, more far-reaching changes would be preferable, though less feasible. For example, some commentators have suggested amending Rule 30 of the Federal Rules of Criminal Procedure, the sole federal rule on criminal jury instructions.⁹³ That rule allows the parties to submit written requests for instructions, but gives courts great leeway in how to handle those requests. Darryl Brown would amend rule 30 to create a “choice-of-instruction rule.”⁹⁴

To his thinking, the normative value of protecting the accused is furthered by requiring the court to give the defendant’s instruction, so long as it is legally sufficient. That view recognizes the decision effects of instruction, and seeks to protect the defendant from adverse effects.⁹⁵ As amended, then, judges “when offered alternative instructions to convey a constitutional rule designed to protect defendants, must assess first whether each version is legally sufficient. If the defendant’s version is legally sufficient, judges must use the instruction as offered by the defendant.”⁹⁶

Though Brown may be right, acceptance is unlikely, given his normative assumptions and nuanced concept of decision effects. More likely to gain acceptance might be a process-oriented rule that required some judicial assessment of jury comprehension. John Cronan also recommended an amendment to Rule 30, but of a more modest nature.⁹⁷ Analogizing to the judge’s role in accepting guilty pleas, he suggested a more active role for trial judges. Under his proposal, a judge would have to address the jury before deliberation to assure its understanding of the relevant law. Thus his amendment would require that:

⁹² *Gacy v. Welborn*, 994 F.2d 305, 314 (7th Cir. 1993) (quoting SUP. CT. COMM. ON JURY INSTRUCTIONS IN CRIM. CASES, ILL. PATTERN JURY INSTRUCTIONS – CRIM. (1987).

⁹³ FED. R. CRIM. P. 30.

⁹⁴ Brown, *supra* note 11, at 76.

⁹⁵ See Brown, *supra* note 11.

⁹⁶ *Id.*

⁹⁷ John P. Cronan, *Is Any of This Making Sense? Reflecting on Guilty Pleas to Aid Criminal Juror Comprehension*, 39 AM. CRIM. L. REV. 1187, 1231 (2002).

Before the jury retires to deliberate, the court must address the jury personally in open court and determine that the jury understands the elements of the crime charged, the presumption of innocence, and the prosecution's burden of proving guilty beyond a reasonable doubt. If any jury expresses confusion at any point prior to returning the verdict, the court must explain the charge until satisfied that the confusion is resolved. Any variances from the procedure required by this rule or errors by the court in offering further clarification of the charge which do not affect substantive rights shall be disregarded.⁹⁸

Consider the effect of such a rule on *Weeks*, for example. There, the jury sat passively as the judge read that confusing instruction on the death penalty. Under this rule, it would have focused on the relevant issues and almost inevitably have considered that issue of aggravating and mitigating factors; the very presence of this active judge would have created a learning atmosphere encouraging questions and clarification. Moreover, in the presence of such a rule, juries would be encouraged to consult the judge during deliberations and the judge, as well, encouraged to play a more instructive role.⁹⁹

Unfortunately, its use doesn't align well with current practice. The trial bench would be leery of departing from its current, safe practice of confining itself to legal correctness, and the appellate bench would also be pressed to deal with this dynamic instructional process. Indeed, even Cronan acknowledged that this procedure "could turn ugly;" he nevertheless saw that as far preferable to "the prospect of those confused jurors rendering criminal verdicts without having their confusion resolved."¹⁰⁰

In fact Cronan, like many others went on to recommend a variety of changes that would aid comprehension.¹⁰¹ Yet, though some of these recommendations are hardly threatening, the plain language movement is gaining most traction, perhaps for its recognition of the opacity of much legal language.

Key to this movement is the recognition of the complete reliance of juries on the language of instructions. Earlier, I posited a Story-based theory of jury behavior, in which coherence was lent by the imposition of some template to the proceedings. We might see that in conflict with a more rigid, cognition-based view, dependent as it is on rules captured in words. But

⁹⁸ *Id.* at 1233.

⁹⁹ Naturally, as a substantive matter, had the judge followed the Brown proposal, the jury would have had a clearer instruction, one not undermining the defendant's interests.

¹⁰⁰ Cronan, *supra* note 97, at 1234.

¹⁰¹ At various points he, like many others discussed here, recommended the following:

1. Plain language instructions;
2. Furnishing jurors with printed copies of instructions;
3. Inviting jurors to ask clarifying questions;
4. Extensive preliminary instructions;
5. Permitting juror note-taking;
6. Providing jurors with "cheat sheets" during trial;
7. Incorporating examples in instructions;
8. Visual assistance to accompany instructions;
9. Allowing jurors to read along while receiving instructions;
10. Choice of instruction rule (similar to Brown's, with serious caveats); and
11. Special instructions on difficult issues.

Id. at 1187.

that's only an apparent conflict.

As Herman and Solan have pointed out, people can think simultaneously in two different ways, in an "associative manner," consistent with the Hastie thesis, and in a rule-based manner, receptive to instructions.¹⁰² Thus, whether instructions either shape the narrative or reinforce rule-based thinking, the research amply demonstrates that improved instructions lead to improved performance.

Some judges, in obviously elitist tones, question the capacity of jurors to understand legal complexities. In the *Gacy* case, Judge Easterbrook opined that jurors are often "simply unable to grasp thoughts unfamiliar to them."¹⁰³ By contrast, in a report of the Federal Judicial Center, Judge Marshall commented that "[t]he principal barrier to effective communication is probably not the inherent complexity of the subject matter, but our inability to put ourselves in the position of those not legally trained."¹⁰⁴

That establishes the key principle of the plain language movement: keep your audience in mind.¹⁰⁵ Throughout this piece, I've mentioned the preoccupation of judges with correctness, to the exclusion of communication. Tiersma and others lament this, noting this myopic fixation on accuracy above all else.¹⁰⁶ Recognizing that "communicating is different from merely speaking or reading to someone," he sets out a series of prescriptions for effective communication to the target audience.¹⁰⁷ That includes such seemingly obvious matters as being concrete, using an understandable vocabulary, employing a logical organization and keeping grammatical constructions simple and straightforward.¹⁰⁸

The potential to confuse is particularly great when legal language is at

¹⁰² Susan N. Herman & Lawrence M. Solan, *Introduction: The Jury in the Twenty-First Century: An Interdisciplinary Conference*, 66 BROOK. L. REV. 971, 981 (2001).

¹⁰³ *Gacy v. Welborn*, 994 F.2d 305, 311.

¹⁰⁴ FED. JUDICIAL CENTER, PATTERN CRIM. JURY INSTRUCTIONS (1988).

¹⁰⁵ PETER M. TIERSMA, COMMUNICATING WITH JURIES: HOW TO DRAFT MORE UNDERSTANDABLE INSTRUCTIONS (Loyola-LA Legal Studies Paper No. 2009-44) (2009), available at <http://ssrn.com/abstract=1507298>.

¹⁰⁶ For example, elsewhere, Tiersma noted that the Massachusetts reasonable doubt instruction is taken verbatim from an 1850 case, commenting that "cases and statutes are written primarily for an audience of lawyers and, thus, have never been intended to be read and understood by the lay public." Peter Tiersma, *The Rocky Road to Legal Reform: Improving the Language of Jury Instructions*, 66 BROOK. L. REV. 1081, 1084 (2001). A law professor, Tiersma was a longtime member of a California task force appointed to make its jury instructions more intelligible. That project was successfully completed several years ago. See PETER M. TIERSMA, REDRAFTING CALIFORNIA'S JURY INSTRUCTIONS (Loyola-LA Legal Studies Paper No. 2009-42) (2009), available at <http://ssrn.com/abstract=1504984>.

¹⁰⁷ TIERSMA, *supra* note 105, at 1.

¹⁰⁸ *Id.* As someone genuinely concerned with educating jurors, Tiersma also notes the deficiencies of a system locked into largely giving oral instructions at the end of the trial: "It would make much more sense to give most of the instructions at the beginning of the trial and to give a brief summary, along with a written copy of the earlier instructions, just before deliberations begin." *Id.* at 3.

odds with general parlance. Tiersma calls such language “legal homonyms,” and they are troubling because “psychological studies have shown that it is very hard to dislodge the ordinary meaning of a word once the meaning is established.”¹⁰⁹ Unfortunately, the law is replete with such homonyms, especially in the homicide area, where terms such as malice aforethought and depraved and malignant heart abound. These terms, often emotionally charged though legal terms of art, resonate with jurors in unforeseen ways. The sad little case involving the killing of Trayvon Martin reveals many of these failures to instruct juries properly.

B. THE TRAYVON MARTIN CASE

On February 26, 2012, George Zimmerman killed Trayvon Martin in Sanford, Florida. Thereafter, in a case that attracted international attention, a 6-person jury acquitted Zimmerman. That much is known. Moreover, Zimmerman claimed self-defense and, though he did not testify, apparently prevailed on that basis. That is also known.

The instructions in that case, meticulously tracking the Florida pattern instruction, contain virtually every flaw discussed by Tiersma and others.¹¹⁰ That jury sat over a painful case with indeterminate facts, involving the killing of a teenage boy. Its only understanding of the law would come from those instructions. However, in the very first page of instructions, the court both injected irrelevant items and betrayed any logical principle of organization.

Judge Nelson started reasonably enough by providing an “introduction to homicide.”¹¹¹ However, she then contrasted excusable homicides and justifiable ones, creating an irrelevant distinction, since the jury would only be concerned with self-defense (justification). Worse, she then went on to explain both at length, rather than first explain the relevant categories of homicide (murder and manslaughter). Placing matters in the wrong order could only confuse a jury, for it is only logical to ask *against what* is one asserting that defense in the first place.

Then, without resorting to any concrete examples, the court repeated the language of second-degree murder verbatim, including its requirement of a “depraved mind without regard for human life” that acts from “ill will, hatred, spite or an evil intent.”¹¹² Failing completely to tailor this to the audience before her, the judge resorted to these “legal homonyms” and other inaccessible legal language. Again, imagining our group of foreign lawyers listening in on this, how surprised they might be at a court simply repeating something written by and for lawyers to non-lawyers who must someone apply that law. But recognize also that by then the jury was not

¹⁰⁹ *Id.* at 7.

¹¹⁰ See Florida v. Zimmerman Final Jury Instructions (2013), available at http://www.flcourts18.org/PDF/Press_Releases/Zimmerman_Final_Jury_Instructions.pdf (last visited January 20, 2014).

¹¹¹ *Id.* Hereafter, all references to the instruction are from that, unpaginated, source.

¹¹² *Id.*

only encumbered by the excess baggage of excuse theory, but also by legal terms that simply could not be applied as stated.

But the flaws are systemic, as Judge Nelson was probably simply fighting to keep her head above water throughout much of this. She followed a familiar (though deeply flawed) template thus far. The trial, however, was singly focused on Zimmerman's justification for killing, as that was the only issue. Thus, we would have hoped that the Florida judiciary would have been prepared on this issue, some 16 months after the killing took place. It was not.

It's hornbook law that self-defense is unavailable to an aggressor, as one cannot exploit a self-generated need to kill.¹¹³ Indeed, also part of Florida law, its pattern instructions cover this issue, stating, "the use of deadly force is not justifiable if you find ... the defendant initially provoked the use of force against himself."¹¹⁴ At the eleventh hour, the prosecution sought this instruction, a move strongly opposed by the defense.¹¹⁵ With scarcely a pause, Judge Nelson announced that the court would not give that instruction.¹¹⁶

It shouldn't work that way, a judge being somewhat bullied into an on-the-spot ruling by the invocation of the notion of "error."¹¹⁷ Unfortunately, a confluence of familiar factors forced that mistake, from the fear of reversal to the use of unhelpful instructions in a system that has given little thought to what juries need. Indeed, arguing that jury instructions matter a great deal, one commentator concluded that losing "the initial aggressor instruction may have been the moment the state lost its case."¹¹⁸

But rehashing the verdict is not the point. Rather, avoiding mistakes is; yet, lacking the tools to legally analyze this tragic human conflict, the jury was painfully unequipped for its task. That is, had it been instructed in a sensible manner, it would have first asked, "Who started it?" and gone on from there. Bottomed as it is on necessity, self-defense law requires the

¹¹³ "One who is not the aggressor in an encounter is justified in using a reasonable amount of force against his adversary when he reasonably believes" it is necessary. WAYNE R. LAFAVE, *PRINCIPLES OF CRIM. LAW* 426 (2d ed. 2010). Dressler embellishes on this, pointing out that a person is an aggressor who even starts a *non-deadly* conflict, and the issue of who the aggressor was "is a matter for the jury to decide, based on a proper instruction on the meaning of the term." JOSHUA DRESSLER, *UNDERSTANDING CRIM. LAW* 224 (6th ed. 2012).

¹¹⁴ Florida v. Zimmerman Final Jury Instructions, *supra* note 110, at instruction 3.6(f) (citing FLA. STAT. § 776.041).

¹¹⁵ Thecount.com, *George Zimmerman Trial Jury Instruction Argument 7.11.13 Pt.8*, YOUTUBE (July 11, 2013) http://www.youtube.com/watch?v=1_UY18DMHV4.

¹¹⁶ *Id.* at approximately the 4:10 point.

¹¹⁷ *Id.* Defense counsel rather transparently told the judge that giving the instruction would be "error," with the desired result. *Id.*

¹¹⁸ Alafair Burke, *What You May Not Know About the Zimmerman Verdict: The Evolution of a Jury Instruction*, THE HUFFINGTON POST, July 15, 2013, http://www.huffingtonpost.com/alafair-burke/george-zimmerman-jury-instructions_b_3596685.html.

actor's conduct to have been necessary at each juncture. Yet, knowing nothing about this, the jury was at a complete loss to reason through to a sensible result. Worse, because of this *culture* of non-concern about the *effect* of instructions on jurors, they were not only deprived of vital instructions, but given some sure to produce confusion and worse.

Much discussion of the case has focused on the so-called "Stand Your Ground" law present in Florida. From 2005-2007, spurred by National Rifle Association lobbying, 27 states have broadened their laws on self-defense to eliminate the requirement for retreat.¹¹⁹ Yet, though never a legal issue in this case, stand your ground sentiment may still have figured prominently in the verdict.

Having refused to instruct the jury on the aggressor issue, Judge Nelson then provided an incomplete and misleading instruction on self-defense. First, she told the jury "A person is justified in using deadly force if he reasonably believes that such force is necessary to prevent imminent death or great bodily harm to himself."¹²⁰ Without the aggressor limitation, that is a false statement of law, as it ignores who was at fault in starting the conflict.

Then, despite the absence of a retreat issue under those facts, she instructed the jury on that "stand your ground" provision:

If George Zimmerman was not engaged in an unlawful activity and was attacked in any place where he had a right to be, he had no duty to retreat and had the right to stand his ground and meet force with force, including deadly force if he reasonably believed that it was necessary to do so to prevent death or great bodily harm to himself or another or to prevent the commission of a forcible felony.¹²¹

In that single sentence, the court probably provided the pivotal, decisive instruction in the case. Yet, though a single sentence, it ran 76 words in length, starting with a puzzling negative ("not engaged in an unlawful activity"), and embodying perhaps 10 key concepts, simply lumped together without logical connections or amplification, with a false economy of space and structure making understanding impossible.¹²²

First, she used the term "unlawful activity," without providing any definition, but likely conveying the notion that Zimmerman, as a neighborhood watch coordinator in a gated community was simply doing his job in pursuing perceived threats. Second, that notion was reinforced by the reference to "any place where he had a right to be." Unversed in the law

¹¹⁹ See DRESSLER, *supra* note 112, at 227.

¹²⁰ See Florida v. Zimmerman Final Jury Instructions, *supra* note 110.

¹²¹ *Id.*

¹²² In that one baffling sentence, Judge Nelson referred to: 1. No unlawful activity (double negative?); 2. Attacked; 3. In protected public space; 4. No duty to retreat; 5. Stand his ground; 6. Force with force; 7. Even deadly force; 8. If reasonably believed necessary; 9. To prevent death or great injury; or 10. Prevent a forcible felony. *Id.* This completely fails any litmus test for comprehensibility and coherence.

and lore surrounding the retreat requirement, the jury would not view a public place as an expansion of the “castle doctrine,” but rather as a place that should be free from threats.¹²³ Then, the use of the term “stand his ground” played into the reasonableness requirement, thus unmistakably creating the vision of an innocent man simply repelling whatever force came his way, regardless of whether he instigated the attack. It would have seemed obvious that Zimmerman, then, “reasonably believed” in the necessity to kill.

Then, the court tethered reasonableness to the attempt to prevent the commission of a “forcible felony.” Again, the court used a crucial term of art without any attempt at definition or clarification, again using language with clear emotional content. Thus, it proffered an irrelevant instruction that had the clear capacity to confuse and mislead. Indeed, its *very presence* fortified the defense *story* of a man simply doing his job in the face of a dangerous, hostile attacker.

Finally, the court instructed the jury that if in its “consideration of the issue of self-defense you have a reasonable doubt on the question of whether George Zimmerman was justified in the use of deadly force, you should find George Zimmerman not guilty.”¹²⁴ I’m not sure I know what that means, whether the court should have instead said that reasonable doubt about his “lack” of justification should result in an acquittal. In any event, creating some of the same problems present in *Martin v. Ohio*, this had the capacity of utterly confound the jury.¹²⁵

Subsequent discussions with a juror who came forward prove this. Anderson Cooper interviewed “Juror B 37” several days after the verdict. The following colloquy appears:

COOPER: Did you feel like you understood the instructions from the judge? Because they were very complex. I mean, reading them, they were tough to follow.

JUROR: Right. That was our problem. It was just so confusing what went with what and what we could apply to what. Because I mean, there was a couple of them in there that wanted to find him guilty of something. And after hours and hours and hours of deliberating over the law and reading it over and over and over again, we decided there’s just no way — no other place to go.

COOPER: Because of the two options you had, second degree murder or manslaughter, you felt neither applied?

JUROR: Right. Because of the heat of the moment *and the Stand Your Ground*. He had a right to defend himself. If he felt threat-

¹²³ The “castle” doctrine represents an exception to the retreat doctrine. “Stand your ground” expands that notion to public spaces. See DRESSLER, *supra* note 112, at 228.

¹²⁴ *Id.*

¹²⁵ See *Martin v. Ohio*, 480 U.S. 208 (1987).

ened that his life was going to be taken away from him or he was going to have bodily harm, he had a right.

COOPER: Even though he got out of the car, followed Trayvon Martin that didn't matter in the deliberations. What mattered was the final seconds, minutes when there was an altercation and whether or not in your mind the most important thing was whether or not George Zimmerman felt his life was in danger?

JUROR: That's how we read the law. That's how we got to the point of everybody being not guilty.¹²⁶

I'm not questioning the result. Rather, we have the spectacle of jurors doing their best, betrayed by a system with no real concern for communication. However, a properly instructed jury might well have rendered the same verdict, as there is nothing inherent in the applicable law that, if explained properly in a true learning environment, is beyond their ken, and the facts here were terribly murky and indistinct.

I say this despite all the bleak claims made by scholars, many repeated here, about the incapacity of juries to perform competently. Many of those claims fail, as restructuring the roles of court and jury can address many, if not most, comprehension issues. But Paul Robinson disagrees.

Long an advocate of code reform,¹²⁷ Robinson seems to undercut that very aspiration:

Juries commonly do not understand the instructions that judges give them. If they do understand the instructions, they frequently are unable to remember them or apply them during jury deliberations. Even if they are able to apply them, they sometimes will not apply them if they do not agree with them. The truth is, the liability rules that juries apply at trial are not those of the code. The governing rules are the commonly incorrect or incomplete jury perception of the instructions, or the jury's own intuition of what justice demands, or a combination of the two.¹²⁸

Failing to account for remedial measures, Robinson's claims are overblown. However, the jury system *is* undermined by one, last factor: The frequent failure of codes to speak intelligibly and their retentionist bias favoring antiquated, incoherent concepts. Eluding all forms of plain English explication, this legal miasma threatens juries and courts alike, and

¹²⁶ Nicole Flatow, *Zimmerman Juror Says Panel Considered Stand Your Ground in Deliberations: 'He Had A Right To Defend Himself'*, THINK PROGRESS, July 15, 2013, <http://thinkprogress.org/justice/2013/07/15/2306631/zimmerman-juror-says-panel-considered-stand-your-ground-he-had-a-right-to-defend-himself/?mobile=nc>.

¹²⁷ See, e.g., Paul H. Robinson, Michael T. Cahill & Usman Mohammad, *The Five Worst (and Five Best) American Criminal Codes*, 95 NW. L. REV. 1 (2000), in which the authors painstakingly examine all fifty-two major American codes from several key perspectives. That very labor indicates an assumption that codes have real-world consequences, heavily dependent on their quality.

¹²⁸ Robinson, *supra* note 12, at 170.

must be eliminated. Good instructions on bad law are the proverbial lipstick on a pig.¹²⁹ That won't do.

VI. THE PROBLEM WITH BAD CRIMINAL CODES

Thus far in this discussion, jury mistakes resulted from a variety of external forces, such as poor instructions, a passive, reversal-obsessed judiciary and the occasional drive of juries to misapply the applicable law. Nothing in the law itself posed any insuperable barriers to communication or proper decisionmaking. But just as “legal homonyms” tend to blur thinking, some formal language itself “tends to reduce comprehension.”¹³⁰ And, that's exacerbated by a vocabulary that is not only unwieldy and inaccessible, but in some cases, simply incoherent. The longstanding failure to effect penal reform left us with a criminal law that was “often archaic, inconsistent, unfair, and unprincipled.”¹³¹

The publication of the Model Penal Code¹³² had a “stunning” impact on American code reform.¹³³ Influencing the codes of over 30 states, it stripped much archaic, pointless language from the law, resulting in vastly more functional, comprehensible statutes. Thus, clear codes are thoroughly congruent with plain English instructions, leading to improved jury comprehension and performance. Unfortunately, the opposite is also true: The persistence of poor codes severely hampers efforts to properly educate juries.

To demonstrate this, Tiersma uses the example of “malice aforethought.” Part of the California code, it underpins its definition of murder. Remarkably, the California statute “explains that malice aforethought can be implied when the circumstances attending the killing show an abandoned and malignant heart. Clearly, almost no ordinary person has any idea what an *abandoned heart* is, and *malignant heart* is only marginally better.”¹³⁴ What, then, happens to this when explained in plain English?

Though the new pattern instruction improved its organization and usage, it still dealt with this vexing term, explaining, “there are two kinds of malice aforethought, express malice and implied malice.”¹³⁵ An irrelevant

¹²⁹ During the 2008 Presidential campaign, Obama famously said: “That's not change. That's just calling something that's the same thing something different. You know you can put lipstick on a pig, but it's still a pig. You know you can wrap an old fish in a piece of paper called change, it's still going to stink after eight years.” Rebecca Sinderbrand, Sasha Johnson & Chris Welch, *'Lipstick on a Pig': Attack on Palin or Common Line?*, CNN, Sept. 10, 2008, <http://www.cnn.com/2008/POLITICS/09/10/campaign.lipstick/>.

¹³⁰ TIERSMA, *supra* note 105, at 3.

¹³¹ DRESSLER, *supra* note 112, at 30.

¹³² MODEL PENAL CODE Pt. I: General Provisions (Official Draft and Revised Comments 1985); Pt. II: Definition of Specific Crimes (Official Draft and Revised Comments 1980).

¹³³ Sanford H. Kadish, *The Model Penal Code's Historical Antecedents*, 19 RUTGERS L.J. 521, 538 (1988).

¹³⁴ TIERSMA, *supra* note 105, at 8.

¹³⁵ *Id.* at 30.

distinction generally, it has the troubling capacity to beleaguer jurors with pointless clutter that can only confuse. That's especially so, since the instruction concludes on the note that "[m]alice aforethought does not require hatred or ill will toward the victim."¹³⁶

Understand, then, what's required of jurors; a strange term has been thrown at them, one involving the worst form of legal homonym, yet one they must apply. However, they are *then* told that malice does not really mean malice. Worse, they are then informed that this term yields to the further distinction of "implied" and "express" malice, with the mind-numbing thought that something we call "malice" can be "implied." Thus, as this sleight of hand unfolds, they are told that malice means something quite different from what they might ordinarily think, yet they must then apply it in that strange, unfamiliar manner. That just strains cognitive capacities (and tolerance) past the breaking point, all because of the existence of an archaic code, and the refusal in the CALRIM committee to part ways with that term.

Perhaps even more troubling, this confusing language plays multiple roles in causing problems for juries. David Crump recalls his experience as a trial attorney during *voir dire* noting that the "effort spent unraveling the meaning of malice aforethought consumed a major part of that time, making it impractical to address other important subjects."¹³⁷ Recognizing that jurors could only absorb so many "foreign concepts at one sitting," he theorized that this term seemed to "elbow out" understanding of other important concepts.¹³⁸ Thus, he concluded that "adjudication by metaphor" ("depraved heart") and double misnomer ("malice aforethought") could only muck up the cognitive process, yielding nothing good.¹³⁹

But though Crump asserts that juries eventually fight their way through to verdicts with some of these terms, others are entirely unmanageable.¹⁴⁰ That includes the most well-known formulation of first-degree murder, killings committed with "premeditation and deliberation." Superficially appealing, that concept quickly becomes completely unruly.¹⁴¹

¹³⁶ *Id.*

¹³⁷ David Crump, "Murder Pennsylvania Style": Comparing Traditional American Homicide Law to the Statutes of Model Penal Code Jurisdictions, 109 W. VA. L. REV. 257, 303 (2007).

¹³⁸ *Id.* Indeed, in the classic film *Amadeus*, the Austrian Emperor complained of too many notes, saying to Mozart: "My dear fellow, there are in fact only so many notes the ear can hear in the course of an evening. I think I'm right in saying that, aren't I, Court Composer?" AMADEUS (Warner Bros. 1984); *From the Movie "Amadeus", TOO MANY NOTES*, <http://www.toomanynotes.com/Amadeus.htm> (last visited Aug. 13, 2013) (excerpt of movie transcript).

¹³⁹ Crump, *supra* note 136, at 305.

¹⁴⁰ He notes "it is usually possible to educate the jury about malice aforethought." *Id.* at 303.

¹⁴¹ It's superficially appealing because we think the studied, acted-upon decision to kill most devalues human life. But Dressler even questions that intuition, noting that though the impulsive killer would only be guilty of second-degree murder, "as a function of

Culpability terms are functional so long as there's an objective referent for mental states. Thus, for example, when we consider whether someone "knew" something, we would look to surrounding circumstances and the actor's conduct. For example, in the *Reed* case,¹⁴² it was at least superficially easy to determine that Leroy Reed knew he possessed the prohibited handgun. However, "premeditation and deliberation" are entirely different. Those terms entirely describe private mental events with no clear observable manifestations.

That partially explains the tortured law in the area, in which courts sometimes speak of the absence of any fixed time for "premeditation," characterizing it as something that can take place in the "twinkling of an eye." Yet, that entirely conflates it with lesser degrees of murder, providing a setting in which juries must make this terrible choice, often involving capital murder, with not one iota of real guidance. As Crump said, that jurisprudence of homicide "seems like a contraption held together by duct tape and bailing wire," an intolerable situation.¹⁴³ Thus, one, final story will illustrate the unacceptable problems caused by poor, outdated codes.

A. *LOUISE WOODWARD: A BAD NANNY, BUT NO MURDERER*

"The larger lesson of the au pair case is...about the damaging effects of an archaic criminal code -- damaging not only to defendants, and victims but to all of us."¹⁴⁴

Louise Woodward, a teenage British au pair in Massachusetts, was convicted of second-degree murder in the shaken-baby death of an infant.¹⁴⁵ Under Massachusetts law, that conviction carried a mandatory life sentence, with eligibility for parole only after 15 years served. Judge Hiller Zobel meted out that sentence, having refused to instruct the jury on the lesser crime of manslaughter.¹⁴⁶ Yet, after the conviction, that self same judge reduced that conviction to one for involuntary manslaughter, resentencing her to time served (279 days). This entire scenario resulted from a striking interplay of terrible laws, all of which remain unchanged to this day.

Second-degree murder exists on a common law basis in Massachusetts, as the statute simply refers to it without definition.¹⁴⁷ Having defined first-degree as a killing with "deliberately premeditated malice afore-

depravity or dangerousness, most people would reverse the results." DRESSLER, *supra* note 112, at 505.

¹⁴² See *supra* text accompanying note 37.

¹⁴³ Crump, *supra* note 136, at 350.

¹⁴⁴ Paul H. Robinson, *State's Weird Law Skewed Au Pair Case*, NEWSDAY, Nov. 17, 1977, at A32.

¹⁴⁵ Commonwealth v. Woodward, 694 N.E.2d 1277 (Mass. 1998).

¹⁴⁶ The defense, over the Commonwealth's objection, only requested the instruction on murder. *Id.* at 1283.

¹⁴⁷ MASS. GEN. LAWS ch. 265, §1 (2011); see *supra* text accompanying note 4.

thought,”¹⁴⁸ the Commonwealth was silent on the culpability basis for other murders. Naturally, that again forces a jury to conclude something that just isn’t true: That the defendant acted “with malice.”

The pattern instructions says:

Malice, for purposes of murder in the second degree, also includes (3) an intent to do an act, which, in the circumstances known to the defendant, a reasonable person would have known created a plain and strong likelihood that death will result. Under this third meaning of malice, you must decide whether, based on what the defendant actually knew at the time he acted, a reasonable person would have recognized that his conduct created a plain and strong likelihood that death would result.¹⁴⁹

Thus, the culpability required for second-degree murder in Massachusetts is negligence, a controversial mental state for *any* criminal liability, let alone liability for a crime carrying a mandatory life sentence.¹⁵⁰ Moreover, unlike some forms of “criminal negligence” that require “gross” or “substantial” negligence, this only requires that risk be plain with a strong likelihood of death.¹⁵¹ Relying on a “loopy judge-as-god provision that keeps the pitiful criminal code from self destructing,” Zobel set about undoing

¹⁴⁸ A term that defies any rational interpretation, as it senselessly commingles the entirely separate concepts of premeditation and deliberation. Unsurprisingly, the Massachusetts code ranked a dismal forty-eighth in Robinson’s ranking. Robinson, *supra* note 126, at 61.

¹⁴⁹ MASS. CONTINUING LEGAL EDUC., MASS. SUPERIOR CT. CRIM. PRAC. JURY INSTRUCTIONS § 2.4 (2003) (instruction on second-degree murder). Understand that the reference to “3” means to the third type of situation that can prove “malice.” Indeed, a recent case compounded this bizarre concept by, in tandem, also reciting the thoroughly useless notions of general and specific intent. Thus it said: “The third prong of malice requires a general intent, unlike the first two prongs of malice, which require a specific intent.” *Commonwealth v. Mountry*, 972 N.E.2d 438, 447 (Mass. 2012).

¹⁵⁰ As *mens rea* literally means guilty mind, it’s hard to ascribe that to someone who was simply empty headed about some risk or other. However, even Hart defended criminal liability based on negligence, theorizing that the very availability of the criminal sanction might cause people to be more vigilant thereafter. See H.L.A Hart, *PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW* (1978). As Hart said: “Threats may not only guide your deliberations—your practical thinking—but may cause you to think...The threat of punishment is something which causes him to exert his faculties, rather than something which enters as a reason for conforming to the law when he is deliberating whether to break it or not.” *Id.* at 134.

However, though the Model Penal Code bases liability on negligence, it requires the actor to flout a “substantial and unjustifiable risk.” MODEL PENAL CODE § 2.02(2)(d) (2012). Moreover, it clearly disfavors negligence, as in the absence of a prescribed mental state in a statute, proof of purpose, knowledge or recklessness is sufficient, but not negligence. MODEL PENAL CODE § 2.02(3) (2012).

¹⁵¹ Thus, not only does the Code version require a “substantial and unjustifiable” risk, but also the actor must at least dimly perceive that risk, as it refers to the “circumstances known to him.” MODEL PENAL CODE § 2.02 (2012) Finally, that conduct must represent a “gross deviation” from the reasonable person standard. MODEL PENAL CODE § 2.02(2)(d) (2012).

this wrong.¹⁵²

Ten days after imposing sentence, Judge Zobel reduced the verdict from murder to involuntary manslaughter, after a hearing on post judgment relief.¹⁵³ Following appeals by both the Commonwealth and the defense, the Supreme Judicial Court of Massachusetts let his decision stand.¹⁵⁴ The law-facts mesh in this case revealed the consummate inadequacy of a Massachusetts law that though unusual, finds many counterparts throughout the country.

In support of his decision, Judge Zobel cited the “circumstances in which Woodward acted” as characterized by “confusion, inexperience, frustration, immaturity and some anger, *but not malice (in a legal sense)*.”¹⁵⁵ Using that troubling homonym, he yoked it together with true human characteristics, almost as if he meant “malice” as it is customarily understood. The court was quick to react,¹⁵⁶ “noting a fine line distinguishes murder based on the third prong of malice, from the lesser offense of involuntary manslaughter.” Comparing murder’s “plain and strong likelihood of death” with manslaughter’s “high degree of likelihood that substantial harm will result to another” it saw obvious overlap and room for confusion.¹⁵⁷ Exactly. In this heart-rending case, a lay jury heard extensive, highly detailed expert opinion,¹⁵⁸ painfully rendering the only available verdict. It did not err; the law did.

This failure is plainly in the law itself. In *Woodward*, the jury studiously applied the law, having ample evidence that she breached her duty of care to the child. But that very predicate for liability was grossly flawed, thus producing that terrible result against which Judge Zobel successfully fought. But it shouldn’t have to work that way.

Of all the impediments to justice recounted here, poor statutes are the most troubling and most resistant to change. Legislatures do not spontaneously take up criminal law reform and, indeed, proposing reform can be costly for members.¹⁵⁹ Moreover, whereas many major legislative issues

¹⁵² Robinson, *supra* note 144, at col. 2.

¹⁵³ Rule 25(b)(2) of the Massachusetts Rules of Criminal Procedure vests authority in the judge to either set aside the verdict and order a new trial, or find the defendant guilty of a lesser included offense. MASS. R. CRIM. P. 25(b)(2).

¹⁵⁴ *Woodward*, 694 N.E.2d, at 1277.

¹⁵⁵ *Id.* at 1286.

¹⁵⁶ *Id.*

¹⁵⁷ *Id.*

¹⁵⁸ Indeed, Patrick Barnes, the key witness for the Commonwealth has since concluded that his assessment was incorrect, and that “shaking was irrelevant in that case.” *Frontline Interview: Dr. Patrick Barnes*, PBS, <http://www.pbs.org/wgbh/pages/frontline/the-child-cases/interviews/patrick-barnes.html> (last visited Aug. 13, 2013).

¹⁵⁹ For example, though felony murder represents an abomination long-since abandoned in Britain (from which it came), any practical legislator would be reluctant to recommend repeal, for fear of seeming soft on crime. The cost in political capital can’t be justified.

compel recognition and action,¹⁶⁰ there's simply no lobby for change in our criminal law and, sadly, those recommending it too often wind up simply "talking to each other ... but they do not appear to be talking to anyone else."¹⁶¹

As the late Bill Stuntz also pointed out, "for most of criminal law, no private intermediaries are well positioned to monitor the law's content."¹⁶² Worse, at least as he saw it, interest groups tend to operate from only one direction, broadening liability rules rather than constricting them when appropriate.¹⁶³ Notions such as "malice" and "specific intent" thrive in such an environment, as they provide the elasticity to maintain these broad liability rules. Thus, though it is simply mad to note that only a "fine line" distinguishes involuntary manslaughter from murder (as the Massachusetts court did), though much of the code law we've seen rests on sand, bad law remains resilient, undermining criminal justice incalculably.

VII. CONCLUSION

Surely the jury system in America is riddled with problems. However, few owe to the basic nature of the system. And, though it has come under fire worldwide, its use has, paradoxically, spread in countries as diverse as Argentina, Japan, Korea, Spain, Russia and Venezuela.¹⁶⁴ Such resilience affirms its status as an important civic institution with "enduring attractions" to many.¹⁶⁵

Indeed, any suggestion about limiting its use would have to confront the question of just what would replace it. Juries represent a cross-section of the population with no ideological bent or political affiliation. Ironically, we should find virtue in the very *inexperience* of its members, as they approach matters with a fresh perspective, unencumbered by experience and repeated frustrations. Thus, the chief criticism lies with their ignorance in the legal arena. In fact, many law professors might argue from their difficulties in getting students to grasp fundamental notions to the conclusion that *of course* juries can't get it, given their brief service.

Whatever virtue that has ignores the lack of viable alternative to jury trials. Issues of comprehension are exacerbated by a system largely unconcerned with its natural objectives. The Virginia Committee that happily recounted the adequacy of its instruction in capital cases should be ashamed of itself.¹⁶⁶ The Court only "approved" the fact that the instruc-

¹⁶⁰ That's especially true when commercial and other monied interests are involved. Forces are at work on all aspects of such issues, thus virtually assuring some sort of legislative or regulatory change.

¹⁶¹ William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 508 (2001).

¹⁶² *Id.* at 529.

¹⁶³ *Id.* at 553.

¹⁶⁴ Hans, *supra* note 9.

¹⁶⁵ *Id.*

¹⁶⁶ *See supra* notes 76 and 85 and accompanying text.

tion met minimal constitutional standards, a conclusion utterly refuted by Garvey's incisive, painstaking analysis.¹⁶⁷ Not to revisit that instruction is callous in the extreme. And that was a death penalty instruction.

Similarly, the instructions in the Trayvon Martin case show the same disregard for communication. The plain English movement emphasizes the critical importance of addressing the target audience, something wholly neglected by such instructions. It emphasizes the necessity for concreteness, logical organization, the use of accessible diction and a careful explanatory style, matters also completely ignored. Is it little wonder, then, that the jury there struggled, with some speaking out on their difficulties?¹⁶⁸

But these matters can be addressed, provided there is sufficient political will to do so. Perhaps exposing these blunders can move the system's players to revisit these issues. Revamping archaic criminal codes certainly takes time. However, the measures advocated here can be accomplished easily and swiftly, requiring little cost and effort. Hopefully, the inroads already made will broaden and continue.

¹⁶⁷ See *supra* notes 75-76.

¹⁶⁸ Another juror has come out to reveal her difficulties with the process, Juror B29. Whatever one might say about her reasoning, she labored on a case in which she was ill-served by the Florida system. See Benjamin Mueller, *Zimmerman Juror: "He Got Away with Murder"*, L.A. TIMES (July 25, 2013), <http://www.latimes.com/news/nation/nationnow/la-na-nn-george-zimmerman-juror-b29-20130725,0,3330582.story>.

RAWLS, POLITICAL LIBERALISM, AND THE FAMILY: A
REPLY TO MATTHEW B. O'BRIEN

Greg Walker*

Institute of Education, University of London

ABSTRACT

Responding to an article in a previous issue from Matthew B. O'Brien on the impermissibility of same-sex marriage, this reply corrects a misinterpretation of Rawls's understanding of political liberalism and a misdirected complaint against the jurisprudence of the U.S. federal courts on civil marriage and other matters. In correcting these interpretations, I seek to demonstrate that a publicly reasonable case for same-sex civil marriage is conceivable in line with political liberalism. I conclude the article by arguing that, although the same-sex civil marriage issue is likely to be a matter of controversy for some time in western societies, a proper understanding of the theoretical issues at stake may contribute to a partial de-escalation of the 'culture wars' currently surrounding the issue.

CONTENTS

I.	INTRODUCTION	38
II.	HOW O'BRIEN DOES NOT GET RAWLS'S POLITICAL LIBERALISM QUITE RIGHT	40
III.	HOW A CASE FOR SAME-SEX CIVIL MARRIAGE CAN BE MADE WITHIN THE BOUNDS OF POLITICAL LIBERALISM	47
IV.	SAME-SEX MARRIAGE, 'RATIONAL BASIS' REVIEW AND U.S. CONSTITUTIONAL LAW.....	59
V.	CONCLUSION	65

I. INTRODUCTION

Matthew B. O'Brien's article in a previous issue¹ is worthy of note in that no other opponent of same-sex civil marriage in the U.S. (hereafter SSCM)² has maintained that the oft-cited arguments for SSCM are *contra-ry* to Rawlsian public reason. While others have suggested that arguments *for* traditional opposite-sex marriage are publicly reasonable none as far as I am aware claim, as O'Brien does, that Rawls's own written statements on the family—which appear at least open to SSCM to most interpreters—can be used directly *against* SSCM. O'Brien's article is forthright, raising matters about the recognition of SSCM that need to be addressed squarely by legal and political theorists. Although O'Brien's arguments are not completely original, he presents us with the most sustained treatment of the 'functional' or 'empirical' purpose of civil marriage with a view to resisting the case for SSCM. The view that natural social reproduction is the central publicly reasonable argument in favor of restricting civil marriage to opposite-sex couples has been argued before by other thinkers.³

Addressing this argument is important in juristic terms as the legal argumentation employed by many opponents of SSCM has refocused on the argument that the restriction of marriage to opposite-sex couples is justified by the procreative and reproductive function of heterosexual mar-

* Graduate student, Institute of Education, University of London, 2013-; Research Associate, Cardiff School of European Studies, Cardiff University, 2011-2012; Ph.D., Political theory, Cardiff University, 2012, BSc (Hons.) Politics, University of Bristol, 1997. Elements of section III of the article were presented at the conference 'Liberalism and Republicanism: Public Policy Implications' at University College London's School of Public Policy in February 2013. I am grateful for the encouragement offered by Frank I. Michelman and Cécile Laborde at various stages in writing the article. My thanks also go to two anonymous reviewers and the Editor whose comments and corrections have improved the article. All errors and infelicities remaining are my own.

¹ Matthew B. O'Brien, *Why Liberal Neutrality Prohibits Same-Sex Marriage: Rawls, Political Liberalism, and the Family*, 1 BRIT. J. AM. LEGAL STUD. 411 (2012).

² I use same-sex *civil* marriage prevalently (rather than the term same-sex marriage) as I take the view that there is clearly a crucial aspect of marriage as a concept that relates to its religious or deeper ethical dimension. Given that I adhere to the principle of religious freedom, civil law should not force any religious institution to hold that any particular law on civil marriage is truly just or that the religious institution should be impelled to solemnise or bless civil marriages that it does not recognise as marriages in a religious or sacramental sense. This is not a point of marginal importance.

³ O'Brien's argument is reminiscent of the Catholic political and ethical philosopher Martin Rhonheimer's internal critique of Rawls's political liberalism on matters such as same-sex marriage. See Martin Rhonheimer, *The Political Ethos of a Constitutional Democracy and the Place of Natural Law in Public Reason: Rawls's Political Liberalism Revisited*, 50 AM. J. JURIS. 41 (2005). The Jewish natural law theorist David Novak has similarly argued that natural social reproduction is the key argument for traditional heterosexual civil marriage. See David Novak, *Response to Martha Nussbaum's--A Right to Marry?*, 98 CAL. L. REV. 51 (2010).

riage.⁴ O'Brien's article may be seen by some as a reflection of this new emphasis from the perspective of legal and political theory as he eschews using arguments rooted in the moral disapproval of same-sex relationships advanced by some Roman Catholic natural law theorists.

Matthew O'Brien boldly seeks to turn the tables on advocates of SSCM by arguing that all arguments for SSCM are based on the *moral approval* of same-sex relationships and are thus out of bounds as arguments within the terms of John Rawls's conception of public reason, as outlined in his seminal later works. O'Brien puts forward sustained arguments that any specifically moral arguments for or against SSCM are not admissible as public reasons in Rawls's schema and so we have to look for purely political reasons to justify the institution of civil marriage. O'Brien's exclusive candidate for a purely political reason for the existence of the institution of civil marriage is the political function served by opposite-sex couples in the way that they foster the natural social reproduction of a political community.⁵ All other arguments—for or against SSCM—are partly or wholly based on moral viewpoints that are not directly relevant to political argumentation in the Rawlsian scheme.⁶

My first point in reply to O'Brien is that he is not successful in his central claim that there are no public reasons for the recognition of SSCM whereas there is, in his view, a clear public reason for the exclusive recognition of opposite-sex marriage. He is not successful because O'Brien omits or glosses important and relevant aspects of Rawls's conception of political liberalism in his presentation of it (see section II of this article) and that he does not properly assess the merits of the case for SSCM. In making a counter-argument, I will synthesize an argument for SSCM that is expressible in the language of public reason, drawing on elements present in the public political culture of the United States in its common law and in more recent U.S. constitutional case law that is not *derived* from a comprehensive philosophical anthropology (section III), though it may be *consistent* with some reasonable comprehensive conceptions of human nature.

This reply will challenge aspects of O'Brien's claim that U.S. courts have misapplied the rational basis test to SSCM cases (section IV) by implicitly adopting aspects of Rawls's political-legal theory into the rational basis review - but in doing so the courts have (according to O'Brien) misunderstood key aspects of Rawls's conception of public reason with problematic consequences. It will be beyond the scope of this reply to address all of the numerous sub-arguments O'Brien marshals to defend the overall thesis of his article.⁷ I will concentrate on the main contours of his argu-

⁴ This has been noted by a number of commentators and was wryly referred to by Chief Judge Vaughn R. Walker in his summary of the Proposition 8 proponents' case in his judgment in *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921, 931 (N.D. Cal. 2010).

⁵ O'Brien, *supra* note 1, at § IV.

⁶ *Id.* at § V.

⁷ If anything O'Brien - by including a range of argumentative 'hostages to fortune' which do not seem to be crucial to his overall philosophical argument and may be interpreted by

ment, which is that there is a clear political value or justification for civil marriage status exclusively for opposite-sex couples but that there is no political value to be gained from recognizing SSCM.

II. How O'Brien Does Not Get Rawls's Political Liberalism Quite Right

The core of O'Brien's argument against SSCM is his contention that all arguments for it fail the test of being admissible in terms of the political and legal theory expounded in John Rawls's highly influential treatise, *Political Liberalism*.⁸ O'Brien supplements this argument by highlighting a passage in Rawls's *Justice as Fairness: A Restatement*⁹ that supports the role of the family in reproducing a political society over time, going beyond Rawls to argue that this is the sole publicly reasonable justification of civil marriage – and that this justification applies only to same-sex couples. In this section I aim to demonstrate that O'Brien misinterprets important elements of *Political Liberalism* in relation to the borderline between the domains of the 'moral' and the 'political' and therefore misconstrues Rawls's notion of public reason. This is important for O'Brien's line of argument because he seeks to represent arguments in favor of SSCM as intrinsically *moral* while Rawls's theory of political liberalism is abstemiously and strictly *political*. As we shall see, the issues involved are not so clear-cut.¹⁰

I also clarify Rawls's recognition of the importance of the family in relation to the social reproduction of a political society, arguing *contra* O'Brien, that Rawls's idea of social reproduction through procreative family life does not preclude the recognition of SSCM and that the passages relevant to this in his oeuvre do not demonstrate a putative Rawlsian backing for O'Brien's view that civil marriage should be available only to opposite-sex partners.

John Rawls, in the works cited, is characteristically nuanced in the way he demarcates the boundaries between the moral/philosophical and

some readers as betraying a particular ideological perspective - does not help the reader make a clear judgment on the two key arguments he presents. These argumentative statements include: that low birth rates, such as in Western Europe, threaten the destabilisation of those societies (*Id.* at 432), that large scale immigration "threaten[s] to undermine [European societies'] public political culture" (*Id.* at 433), that people without children are less concerned about intergenerational justice (*Id.* at 435), that contraception is generally ineffective and unreliable (*Id.* at 440-41), that the Association of American Psychologists persists in using discredited methodology in its analyses of same-sex parenting (*Id.* at 444), and that parents who have children through gamete donation act immorally and unjustly to their future children (*Id.* at 441-48).

⁸ JOHN RAWLS, *POLITICAL LIBERALISM* (2005).

⁹ JOHN RAWLS, *JUSTICE AS FAIRNESS: A RESTATEMENT* (2001).

¹⁰ That O'Brien operates with these hard and fast distinctions in his treatment of the SSCM issue is confusing because at one point he acknowledges that Rawls's political liberalism involves moral ideas, though he seems to think that this is limited "to the moral idea of equal citizenship". See O'Brien, *supra* note 1, at 424.

the political. Rawls explicitly writes of a political conception of justice being “of course, a moral conception, it is a moral conception worked out for a specific kind of subject” that is for the “basic structure of a constitutional democratic regime”.¹¹ Rawls therefore does indeed hold that we should view the domain of the political as being *distinct* from other facets of our lives, but at the same time he recognizes that a citizens’ moral values cannot be seen as *separate* from, or in conflict with, fundamentally political values.¹² Inversely for Rawls, “[p]olitical conceptions of justice are themselves intrinsically moral ideas, as I have stressed from the outset. As such they are a kind of normative value.”¹³

This distinction has been clearly recognized by interpreters of Rawls such as the prominent political theorist Gerald Gaus, who states in relation to Rawls’s theory:

[t]hat a belief is moral, religious or philosophical does not itself show that it is comprehensive or general. Indeed, Rawls himself indicates that the political conception has moral, epistemological and metaphysical elements. Moral, religious and philosophical beliefs need not be, and very often are not, comprehensive or general.¹⁴

This is a common enough interpretation of the connections between the moral and the political in *Political Liberalism*,¹⁵ though it is not an interpretation mentioned by O’Brien. He may, however, respond that interpretations of Rawls’s work do, of course, vary.¹⁶

In trying to portray *Political Liberalism* in a strictly neutralist light O’Brien seizes on the phrase “purely political”¹⁷ used by Rawls to describe an aspect of his theory. Although Rawls does use this construction a handful of times in his essay ‘The Idea of Public Reason Revisited’¹⁸ in outlining his own normative view of political liberalism, when he does so he is careful to add the immediate qualification “although political values are intrinsically moral.”¹⁹ As we have seen, Rawls is upfront about the fact that he conceives political liberalism to be a moral conception of human life and value, save that the form and subject matter of that moral content is directed to the basic structure of a political society and is therefore *partial*

¹¹ RAWLS, *supra* note 8, at 175.

¹² “Nor does it [political liberalism] say that political values are separate from, or discontinuous with, other values”. RAWLS, *supra* note 8, at 10.

¹³ *Id.* at 484 n.91.

¹⁴ Gerald F. Gaus, *Reasonable Pluralism and the Domain of the Political*, 42 INQUIRY: INTERDISCIPLINARY J. PHIL. 259, 263 (1999).

¹⁵ See, for instance, Martha Nussbaum, *Political Liberalism: A Reassessment*, 24 RATIO JURIS 1 (2011); Martha Nussbaum, *Political Objectivity*, 32 NEW LITERARY HIST. 887, 891-94 (2001).

¹⁶ For a brief survey of interpretations, see Anthony Simon Laden, *The House That Jack Built: Thirty Years of Reading Rawls*, 113 ETHICS 367 (2003).

¹⁷ See O’Brien, *supra* note 1, at 7, 11, 15.

¹⁸ See RAWLS, *supra* note 8, at 457, 461, 486.

¹⁹ *Id.* at 446 n.19.

rather than comprehensive.²⁰

Reading *Political Liberalism* as advancing a strict form of liberal moral neutrality is not helpful or fully accurate. In contrast Peter de Marneffe sets out a helpful typology between liberal neutrality and perfectionism that helps us see that John Rawls's later political theory is not strictly neutralist but is a form of what de Marneffe calls “deontological perfectionism”. He takes this view because, for de Marneffe, “Rawls clearly holds that it is wrong for the government to limit *basic* liberties for the reason that exercising them is base or unworthy, [but] he apparently allows *non-basic* liberties to be limited for this reason, and confines the basic liberties to those that are “truly essential.”²¹ Rawls himself writes strikingly that neutralism beyond the strict bounds of basic justice and constitutional essentials is “neither attainable nor desirable.”²²

Rawls thus denied being a strong neutralist—and held that the term neutrality itself was “unfortunate” when used in connection with his theory of political liberalism.²³ He resisted the notion of neutrality as put forward by other liberal theorists such as Ronald Dworkin and, though he conceded that his approach did have a certain “neutrality of aim”, Rawls did not believe that his theory was consistent with a “neutrality of effect” as other liberals had proposed.²⁴ In fact O’Brien’s entire narrative in relation to liberal moral neutralism would arguably have been better applied to the early work of Ronald Dworkin (in particular his widely cited essay ‘Liberalism’)²⁵ than the later work of John Rawls. In this regard Jonathan Quong rightly distinguishes between Rawls’s *political* antiperfectionism and Dworkin’s comprehensive (i.e. moral *and* political) antiperfectionism.²⁶ This is not a marginal reading of Rawls, as a varied range of interpreters other than those cited stress aspects of Rawls’s works that do not easily fit with O’Brien’s view of Rawls as a strict moral neutralist.²⁷

²⁰ Rawls writes that, notwithstanding the priority of the right over the good, this “does not mean that ideas of the good need to be avoided: that is impossible”. See RAWLS, *supra* note 8, at 203.

²¹ See Peter de Marneffe, *Liberalism and Perfectionism*, 43 AM. J. JURIS. 101 (1998).

²² RAWLS, *supra* note 9, at 91 n.13.

²³ Rawls writes that he uses the term neutrality as a “stage piece” and as a way of *contrasting* his position with other liberal theories. See RAWLS, *supra* note 8, at 191. Although O’Brien (*supra* note 1, at 424) notes that Rawls did not use the “idiom” of neutrality, O’Brien nonetheless breezily proceeds to label Rawls as a moral neutralist numerous times in his article.

²⁴ See RAWLS, *supra* note 8, at 193, 194.

²⁵ Ronald Dworkin, *Liberalism*, in PUBLIC AND PRIVATE MORALITY 191 (Stuart Hampshire ed., 1978).

²⁶ JONATHAN QUONG, LIBERALISM WITHOUT PERFECTION 21-21 (2010).

²⁷ Steven Wall, for instance, goes as far as interpreting a key aspect of Rawls’s original theory of justice (‘the Aristotelian Principle’ in the unduly neglected Part III of Rawls’s *A Theory of Justice*) as being, in a sense, perfectionist. See Steven Wall, *Rawlsian Perfectionism*, 10 J. MORAL PHIL. 573 (2013).

One of O'Brien's tactics in disregarding the work of advocates of SSCM is to select quotes from liberal authors in favor of SSCM that would appear to confirm that they are 'comprehensive' liberals—in that they may, for instance, value moral autonomy over other conceptions of human agency. He then infers that some or all of the cited advocates of SSCM are comprehensive liberals and hence their arguments in favor of SSCM are, by that very fact, ruled out of bounds because comprehensive liberalism is not consistent with Rawlsian political liberalism.²⁸

What this neglects to take into account is that Rawls clearly does not consider that advancing reasonable comprehensive reasons for supporting certain political measures relating to the basic structure disqualifies the citizen from participating in public reasoning. Following criticism of the first edition of *Political Liberalism*, Rawls clarified his position on public reason in his article 'The Idea of Public Reason Revisited'. There he outlined a 'Proviso': that "comprehensive doctrines, religious or non-religious, may be introduced into public political discussion at any time, provided that in due course proper political reasons.... are presented" in relation to the matter under discussion.²⁹ In other words people may freely mix and concurrently advance both public and comprehensive reasons for measures relating to the basic structure and be considered responsible and public-spirited citizens.

O'Brien in contrast fails to consider the 'Proviso' and writes, incorrectly, that "what Rawls prescribes citizens in a pluralistic democracy should do [is to]: filter their comprehensive doctrines through the deliberative screen of public reason before proposing grounds for legislation."³⁰ As we have just seen, Rawls proposes no such filtering process, only a stipulation that comprehensive moral or philosophical reasons should not be advanced without *any* subsequent public reasons. This clear misinterpretation may explain why O'Brien supposes that comprehensive liberals somehow disqualify their publicly reasonable arguments for SSCM when they venture their own more comprehensive (and contestable) views about autonomy or the *ultimate* nature or purpose of marriage or of human sexuality generally.

This somewhat more permissive approach to public reasoning works with Rawls's wider theory because he allows persons with a comprehensive worldview to be considered *publicly* reasonable (as well as rational) if they can fit a liberal political conception of justice as a 'module' within their comprehensive *weltanschauung*.³¹ Again, O'Brien does not mention this important aspect of Rawls's conception of public reason. Such a 'module' in Rawls's formulation has its own internal principles and reasons that may be consonant with a wider metaphysical or moral comprehensive doctrine,

²⁸ O'Brien, *supra* note 1, at 434, 454-62.

²⁹ See John Rawls, *The Idea of Public Reason Revisited*, 64 U. CHI. L. REV. 783-84 (1997).

³⁰ O'Brien, *supra* note 1, at 450.

³¹ RAWLS, *supra* note 8, at 145.

but not be immediately derived from the comprehensive doctrine held by the citizen. A political conception of justice would be freestanding in the sense that it *includes* “no specific religious, metaphysical or epistemological doctrine beyond what is implied by the political conception itself”³² though, as we have seen, it should still be seen as a partial moral conception as it necessarily includes certain goods of citizens which (as O’Brien himself notes) relate to certain fundamental human needs.

The ‘Proviso’ and the notion of a ‘module’ within political liberalism renders O’Brien’s claim that “an argument [made by Stephen Macedo for SSCM] fails *because* it relies on the assumption that homosexual sexual relationships are intrinsically valuable” highly questionable.³³ An argument does not ‘fail’ for Rawls—I presume in the sense of being impermissible in public reasoning—because it includes or refers to moral claims from a reasonable comprehensive doctrine. Such arguments *are* genuinely permitted in the idea of public reason as long as those moral arguments are, at some point, accompanied by arguments involving distinctively political values.³⁴ This is indeed exactly what Macedo provides in the article cited by O’Brien and there are no grounds for O’Brien to claim that arguments for the *political* value of SSCM (in terms of equal civil rights or primary goods) are somehow disqualified by the fact that this or that theorist may also advance substantively *moral* arguments for treating committed same-sex relationships with respect, or hold more generally that people should be considered morally autonomous.

O’Brien compounds his partial misinterpretation of Rawls³⁵ by appearing to apply a double standard to the arguments put forward by proponents and opponents of SSCM. He is dismissive of those like Macedo and others who may simultaneously advance both public and (in some ways) reasonable comprehensive reasons for recognizing SSCM, writing that it is “the case *in favor* of same-sex marriage that has impermissible

³² *Id.* at 144.

³³ See O’Brien, *supra* note 1, at 428 (emphasis added). O’Brien is referring to an argument Stephen Macedo put forward for SSCM and gay rights generally. O’Brien generally accuses Rawlsians of relying “illicitly on their comprehensive religious or secular doctrines about ‘liberated’ sexual morality in order to single out homosexual relationships *as such* for special promotion”, *supra* note 1, at 437. Macedo, in fact, argues that committed, long term relationships – whether same-sex or opposite sex – are intrinsically valuable and he does not single out same-sex relationships as having any ‘special’ value. See Stephen Macedo, *Sexuality and Liberty: Making Room for Nature and Tradition?*, in *SEX, PREFERENCE AND FAMILY* (David M. Estlund & Martha Nussbaum eds., 1998).

³⁴ Macedo, in the essay referred to by O’Brien, also advances arguments clearly relating to political values when he gives reasons for SSCM on the basis of “social welfare” and the various demonstrable public health and other beneficial externalities of committed relationships (including same-sex relationships). See Macedo, *supra* note 33, at 92-94.

³⁵ I write ‘partial’ here because there is much in O’Brien’s presentation of Rawls’s theory that is both comprehensive and fair, which is why it is surprising that he manages to mischaracterize Rawls in the important ways that I point out in this reply.

motivations that are fatal to legislation”.³⁶ Yet O’Brien inexplicably permits *opponents* of SSCM to use both comprehensive and public reasons for resisting SSCM, without this having any adverse impact in terms of their admissibility in public reason. Referring to the ‘traditional marriage movement’, he concedes that “[m]uch of this movement deploys specifically religious arguments in its defense, but this fact is irrelevant so long as some of these arguments can be re-stated in terms of public reasons...”.³⁷ I cannot understand why O’Brien does not apply the same standard to both sides of the debate, which would surely be the Rawlsian approach. Advocates and opponents of SSCM should both be viewed as reasoning publicly if they use arguments from a comprehensive worldview *and* public arguments centered on political values, or just the latter.

As we have seen, political liberalism as a theory of political principles does not divorce political value from the human good but it does seek to focus the role of a political society on promoting those goods that are relevant to persons as citizens. This is clear from Rawls’s conception of primary goods, which he develops from his initial treatment of them in *A Theory of Justice*,³⁸ as O’Brien notes.³⁹ The primary goods are those goods that any *citizen* would reasonably seek regardless of whatever else they sought. They are of a broad scope and include basic civil and political rights, the “social bases of self-respect” and “income and wealth.”⁴⁰ These primary goods allow citizens a sense of self-worth that enables them to pursue a plan of life.⁴¹

One key feature of justice as fairness – Rawls’s own proposal for a liberal conception of political justice – is that “it is constructed on the basis of the shared fundamental ideas implicit in the public political culture [of a democracy] in the hope of developing from them a political conception of justice.”⁴² This is a point not explicitly mentioned by O’Brien. This shared public political culture “comprises the political institutions of a constitu-

³⁶ O’Brien, *supra* note 1, at 460.

³⁷ *Id.* at 449. O’Brien explains this further when he writes that “[t]he reliance of Rev. [Martin Luther] King [Jr.] and others upon the controversial comprehensive doctrines of the Christian moral tradition did not violate the canons of public reason, however, because the case for racial equality could be re-stated in non-sectarian terms that expressed a purely political conception of justice.” See O’Brien, *supra* note 1, at 449. These statements are consistent with the Rawlsian *Proviso* just outlined, which is not referred to by O’Brien, and I agree with O’Brien in the case of Dr King.

³⁸ JOHN RAWLS, *A THEORY OF JUSTICE* chapter II, § 15 (1971).

³⁹ O’Brien, *supra* note 1, at 425.

⁴⁰ Though Rawls wrote at times of his own sympathies with the idea of human capabilities (as outlined by Amartya Sen and others) when he affirmed that that “basic capabilities are of first importance and that the use of primary goods is always to be assessed in the light of those assumptions”, in this passage he appears to have held that the capabilities should come into play as a constructivist throughput from, rather than an ethical input into the original position. See RAWLS, *supra* note 8, at 183.

⁴¹ See RAWLS, *supra* note 9, at 58-59.

⁴² See RAWLS, *supra* note 8, at 100-101.

tional regime and the public traditions of their interpretation (including those of the judiciary), as well as historic texts and documents that are common knowledge.”⁴³

This is what other theorists categorize as a form of *secondary* constructivism, or *political* constructivism as Rawls himself described it. It is a secondary form of constructivism because it takes as its starting point a thin moral psychology⁴⁴ and the pre-existing content of a public political culture coupled with an understanding of citizens as free and equal. Rawls does not subject these primary presuppositions to a strict procedure of moral construction from the bare minimum of human rationality alone - as some more stringently constructivist theories do.⁴⁵ It is therefore for good reason that readers of Rawls’s political theory even interpret him as including *ethical* elements in his political theory via the notion of the public political culture.⁴⁶

It is worth noting that ‘justice as fairness’, Rawls own version of a liberal political conception of justice, was never intended by Rawls to be the only possible liberal conception. Rawls set out certain broad characteristics of any political conception of justice that may be considered a *liberal* conception, which include:

certain basic rights, liberties, and opportunities (of the kind familiar from constitutional democratic regimes); second, it assigns a special priority to these rights, liberties and opportunities, especially with respect to claims of the general good and of perfectionist values; and third, it affirms measures assuring all citizens adequate all-purpose means to make effective use of their basic liberties and opportunities.⁴⁷

As I see it O’Brien elides Rawls’s understanding into his own by reading Rawls as holding that social reproduction is the *only* reason for marriage, whereas in my judgment Rawls’s position is, at the very least, open to the idea of SSCM. This is not to say that anyone committed to political liberalism of a Rawlsian hue should *ipso facto* be convinced of the case for

⁴³ *Id.* at 13-14.

⁴⁴ The moral psychology predicated in political liberalism is noted by O’Brien, *supra* note 1, at 424.

⁴⁵ For a fuller exploration of the important differences between primary and secondary constructivism in Rawls and other theorists, see PERI ROBERTS, *POLITICAL CONSTRUCTIVISM* (2007).

⁴⁶ See James Gordon Finlayson & Fabian Freyenhagen, *The Habermas-Rawls Dispute: Analysis and Reevaluation*, in *HABERMAS AND RAWLS: DISPUTING THE POLITICAL* 15 (James Gordon Finlayson & Fabian Freyenhagen eds., 2011) (“Political values and ideas taken from the public political culture [and inputted into the constructivist procedure in Rawls’s theory] might include materials that Habermas would classify as ethical rather than moral.”). In Habermas’s political theory ‘morality’ refers broadly to intersubjective social and moral norms governing the common life of a community (or humanity as a whole), whereas ‘ethics’ refers to an individual’s own understanding of ‘the good life’ or personal fulfillment, which can include substantive religious or metaphysical dimensions.

⁴⁷ See RAWLS, *supra* note 8, at 223.

SSCM. That is, after all, a product of deliberative public reasoning and personal judgment.

Quoting Rawls on the family in his later writings, O'Brien writes:

Indeed, Rawls emphasizes that in principle, "[n]o particular form of the family (monogamous, heterosexual, or otherwise) is so far required by a political conception of justice so long as it is arranged to fulfill these tasks [of social reproduction] effectively and does not run afoul of other political values" That is, for political liberalism the state interest in the family is purely functional, even if families in their own self-image are not, and so there is no antecedent political preference for either "traditional" or "liberated" family forms as such.⁴⁸

O'Brien's comment misses the significance of a key point in Rawls's statement he quotes, that such politically recognized forms should "not run afoul of other political values" which could of course refer to those legally approved family forms being in accordance with norms of justice, not breaching citizens' civil rights, or failing to provide for the primary goods of citizens.

The next section of this article aims to demonstrate that a publicly reasonable argument can be made for SSCM along Rawlsian lines, with support from aspects of the public political culture of the United States.

III. HOW A CASE FOR SAME-SEX CIVIL MARRIAGE CAN BE MADE WITHIN THE BOUNDS OF POLITICAL LIBERALISM

A case can be made for SSCM that is consistent with Rawls's conception of political liberalism on one or both of two grounds: first, that civil marriage is a civil right that should be granted regardless of gender or sexual orientation and that such a right has priority over perfectionist arguments about what constitutes the deepest truth regarding marriage derived from a metaphysical or theological anthropology. (This claim requires further argumentation, as O'Brien rightly notes, as to the nature and purpose of marriage from the perspective of civil law).⁴⁹ Secondly, it can be argued that a bar on SSCM is a denial of citizens' access to what Rawls names primary goods, which are themselves expressions of basic human needs, particularly in relation to the primary good Rawls calls the 'social bases of self-respect'.⁵⁰

The civil rights-based and primary goods argument for SSCM to a certain degree overlap, as justifications for either depend on whether any discrimination between opposite-sex couples and same-sex couples in relation to access to civil marriage is considered unjust or merely reflects the essential nature and purpose of civil marriage itself (as SSCM opponents argue). I argue that the relationship between committed and loving same-sex cou-

⁴⁸ See O'Brien, *supra* note 1, at 433-34.

⁴⁹ O'Brien, *supra* note 1, at 421.

⁵⁰ RAWLS, *supra* note 9, at 59.

ples exhibits certain essential features that are present in committed opposite-sex couples to the degree that the denial of same-sex civil marriage rights constitutes arbitrary treatment which fails to ‘treat like cases alike’ without good reason. Therefore the denial of SSCM in the United States can be considered a breach of fundamental justice (rendering another ‘his/her due’ – ‘*suum cuique*’). Such an argument does not depend on metaphysical argumentation inaccessible to public reason and can be seen to flow from reasoning about basic justice consistent with Rawls’s notion of narrow reflective equilibrium.⁵¹

In this section I argue that there is a sufficient basis in the *public political culture* (in the Rawlsian sense) of the United States, expressed through the case law developed from the common law understanding of civil marriage, to justify the assertion that access to both opposite and same-sex marriage should be considered a civil right within a properly *liberal* political conception of justice (in Rawlsian terms).⁵² The denial of SSCM might also constitute the undue withholding of a primary good—the social bases of self-respect—to a fellow citizen in a way that does not respect their fundamental dignity. I make this claim on the basis of the evolution of its juridical understanding of marriage from its source in the western Christian understanding of marriage through to a more recent understanding, which properly distinguishes the procreative understanding of marriage from the other goods and purposes that civil marriage enables.

In setting out a publicly reasonable case for SSCM from the public political culture, I do so while taking up the challenge of Robert George and his collaborators that a political society must come to at least some minimal determination of what the essence of *civil* marriage is before it can adequately respond to the issue of SSCM.⁵³ This minimal determination is unavoidable in some respects,⁵⁴ though any definition must of course be careful not to derive its content from a controversial metaphysical anthropology that could not be shared by adherents of a range of different reasonable comprehensive doctrines (in Rawls’s nomenclature).

I would briefly define *civil* marriage, from the perspective of the legal official (the internal point of view),⁵⁵ as the government’s rightful recognition of the formation of a loving union (of an indefinite term) comprising a

⁵¹ See RAWLS, *supra* note 9, at 30-32 for more on his understanding of narrow and wide reflective equilibrium, concepts that Rawls was instrumental in developing in ethical and political theory.

⁵² In the sense described by Rawls at the close of § II of this article.

⁵³ Sherif Girgis, Robert P. George & Ryan T. Anderson, *What is Marriage?* 34 HARV. J. L. & PUB. POL’Y 248-49 (2010). O’Brien similarly notes – and here we are in agreement – that some definition of civil marriage is needed to avoid question-begging and undue rhetorical flourishes in relation to SSCM, though he does not consider what a possible sex-neutral definition of civil marriage might be; see O’Brien *supra* note 1, at 456.

⁵⁴ For how else might we distinguish a marital relationship from any other type of friendship or personal (non-corporate) association?

⁵⁵ Here I use the internal/external distinction used in H.L.A HART, *THE CONCEPT OF LAW* 57-58, 89-91 (1997).

new human society of two eligible⁵⁶ consenting adults who publicly commit to certain responsibilities, including fidelity and mutual aid, who thus become beneficiaries of certain legal rights.⁵⁷ In this new society the capacity and human need for companionate and affective domestic association helps enable continent sexual expression and can serve a wider good of the orderly reproduction of a political society (and the human race generally). These are the common goods proper to the conjugal society. (This is a definition for the purpose of the present argument—clearly any developed understanding of civil marriage would have to comprehensively examine the essence of marriage, something I cannot do here.)

Certain features from this definition are of course familiar from the threefold traditional ‘goods of marriage’ formed from the traditional western Christian tradition⁵⁸—each of which is familiar from the Anglican 1662 Prayer Book preface to the marriage service.⁵⁹ These goods are, in their established order: *procreatio* (procreation), *mutuum auditorium* (mutual aid) and the *remedium concupiscentium* (as a remedy for concupiscence, or as a way or legitimately ordering sexual desire).⁶⁰ This tradition is the product of many influences and stages of development, from Augustine’s engagement with Jovinian and St Jerome’s debate on the nature of matrimony, through to the scholastic development of the concept with Thomas Aquinas, all the way to the twentieth century personalist emphasis on the relational and loving character of the marital bond.⁶¹

The influence of this tradition on western legal systems through canon law and, in turn, the common law is well documented. This is the case in

⁵⁶ “Eligibility” referring here to consanguinity rules on the basis that there should be no confusion between essentially familial relationships and conjugal relations.

⁵⁷ As Leslie Green argues, civil marriage is the recognition of what already “exists as a matter of social or religious practice” – giving a de jure recognition to a de facto union. See Leslie Green, *Sex–Neutral Marriage*, 64 CURRENT LEG. PROBS 1, 7-8 (2011).

⁵⁸ A clear and concise overview on this development is John Witte, Jr., *The Goods and Goals of Marriage*, 76 NOTRE DAME L. REV. 1019 (2001).

⁵⁹ The words are “[f]irst, It [i.e. marriage] was ordained for the procreation of children, to be brought up in the fear and nurture of the Lord, and to the praise of his holy Name. Secondly, It was ordained for a remedy against sin, and to avoid fornication; that such persons as have not the gift of continency might marry, and keep themselves undefiled members of Christ’s body. Thirdly, it was ordained for the mutual society, help, and comfort, that the one ought to have of the other, both in prosperity and adversity. Into which holy estate these two persons present come now to be joined.” *The Solemnization of Matrimony*, in THE BOOK OF COMMON PRAYER (1662).

⁶⁰ Augustine wrote of the three aspects of marriage as being *proles* (children), *fidelium* (fidelity), and *sacramentum* (symbolic stability). See Witte, *supra* note 58, at 1030.

⁶¹ As seen in Sections 47 to 50 of the Second Vatican Council’s pastoral constitution, *Gaudium et Spes*, Dec. 7, 1965: AAS58 1025-1115, §47-50 (1966), which strongly foregrounds the centrality of conjugal love in its presentation of marriage. See *Pastoral Constitution on the Church in the Modern World: Gaudium et Spes*, THE HOLY SEE (last visited Nov. 25, 2013) (announced December 7, 1965, and published in 1966), http://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_const_19651207_gaudium-et-spes_en.html.

the U.S. states, which inherited the common law and canon law understanding of marriage from England from the pre-revolutionary era.⁶² This understanding evolved over time according to the peculiarities of each state's law, with constitutional provisions being invoked at the federal level from time to time when matters of great import arose. Procreation and the other goods of marriage were often cited in the case law of the U.S. courts.⁶³

In the Catholic understanding of this tradition, though there is more than one 'end' or 'aim' to marriage "[t]hese aims can....only be realized in practice as a single complex aim" and, despite the traditional ordering of the ends of marriage, "there is no question of opposing love to procreation nor yet of suggesting that procreation takes precedence over [conjugal] love".⁶⁴ Indeed the traditional western twofold or threefold ends of marriage,⁶⁵ which do not refer explicitly to love, should not be understood as in any way implying that the foundation of marriage within this understanding is not *love* itself (a communion of love in the original theological language).⁶⁶ What is distinctive in this Catholic tradition—in which the western tradition is rooted—is the notion that the procreative and other ends of marriage are *inseparable* and that what makes a conjugal community different from other forms of human society is its reproductive end, as that tradition sees it.⁶⁷

⁶² See Charles J. Reid, Jr., *The Augustinian Goods of Marriage: The Disappearing Cornerstone of the American Law of Marriage*, 18 BYU J. PUB. L. 449 (2003-2004); Charles J. Reid, Jr., *The Gingerbread Man Thirty Years On: The Parlous State of Marital Theory*, 1 U. ST. THOMAS L.J. 656 (2003); Charles J. Reid, Jr., *Marriage in its Procreative Dimension: The Meaning of the Institution of Marriage throughout the Ages*, 6 U. ST. THOMAS L.J. 454 (2008-2009). Incidentally Professor Reid, who was an outspoken (Roman Catholic) opponent of SSCM when writing these cited articles, in 2013 publicly declared that he has changed his mind and that he now favors same-sex civil marriage.

⁶³ See *id.* (collecting Reid's articles).

⁶⁴ KAROL WOJTYLA, LOVE AND RESPONSIBILITY 68 (1981). Wojtyla later became Pope John Paul II.

⁶⁵ The 1917 Code of Canon Law of the Catholic Church combined the second and third ends of marriage, something perpetuated in later treatments of the ends of marriage in canon law and in *Gaudium et Spes*. See *supra* note 61.

⁶⁶ For one prominent and mainstream Catholic thinker interpreting the magisterial tradition, Martin Rhonheimer, conjugal love is not an abstract love for an end that may or may not issue from the relationship (children) but is a "love for a concrete person" (the spouse), and that "the" "purpose" or "end" of love is the person himself [the spouse], and nothing else". Rhonheimer writes "[l]ove could not be an end to marriage it is rather *its foundation and content*, content that is nevertheless characterised by a natural end in a specific way: at the service of life [referring to the procreative end of marriage]" (emphasis added). See MARTIN RHONHEIMER, ETHICS OF PROCREATION AND THE DEFENSE OF HUMAN LIFE, 86, 87 (2010).

⁶⁷ Martin Rhonheimer calls this strong thesis about the intrinsic bond between sex and procreation, the Catholic 'Inseparability Principle', and this is reflected in the Catholic opposition to all forms of intentional contraception. See Rhonheimer, *supra* note 66, at 44-46, 71-90. I pay particular attention to the Roman Catholic understanding of marriage

An alternative approach is to retain this basic outline of the nature of marriage as a union of love between spouses forming a new society that yields a nexus of human goods for its participants. But in reasserting this basic scheme one can hold that the procreative and mutual aid (companionate) ends of civil marriage are not held to be utterly inseparable and thus incapable of justifying the conferral of civil marriage rights on the basis of the latter (companionate) end of marriage alone.⁶⁸ Being separable as goods or ends does not mean that they cannot be mutually reinforcing as ends in those couples who can procreate. Marriage is understood as a loving union of an indefinite term that produces goods that provide benefit both to the couple concerned and to the wider society.

The possibility of (opposite-sex) civil marriages in which procreation does not occur—or is strictly controlled—has come into sharp relief with the advent of effective contraception in Western societies since the 1960s. This has been reinforced with the increasing phenomenon of new marriages involving post-menopausal women, in part due to the substantial increase in life expectancy. This has inevitably raised the question of how the traditional Western ends of marriage (procreation, mutual aid etc.) can be seen to be absolutely inseparable.⁶⁹

The philosopher G.E.M. Anscombe presciently considered that an acceptance of the principle of intentional contraception (which she opposed) into societal mores could have wider consequences on how civil marriage was conceived. She predicted that the acceptance of sexual expression between partners not founded on a marital relationship intentionally aimed at procreation would create a logic for opening up civil marriage to members of the same-sex, something she clearly opposed.⁷⁰ In other words if the

because many of the most ardent (and philosophically sophisticated) supporters of traditional civil marriage, such as Rhonheimer, John Finnis, Robert P. George and Maggie Gallagher, are Roman Catholic. I surmise that Matthew O'Brien also writes from within this tradition although his arguments differ in part from those advanced by Finnis and George.

⁶⁸ Here I differ with Ralph Wedgwood, who goes too far by altogether omitting procreation within his threefold understanding of the essence of marriage, which for him includes "(1) sexual intimacy; (2) domestic and economic cooperation; and (3) a voluntary mutual commitment". See Ralph Wedgwood, *The Fundamental Argument for Same-Sex Marriage*, 7 J. POL. PHIL. 229 (1999).

⁶⁹ It is worth noting that the Catholic Church has always allowed people who know they are infertile to marry. Cf. CODE OF CANON LAW OF THE CATHOLIC CHURCH c.1084.3 (1983) affirming that "[s]terility neither prohibits nor invalidates marriage". Appreciating the suffering of an infertile couple, CATECHISM OF THE CATHOLIC CHURCH c.1654 (1999) states: "Spouses to whom God has not granted children can nevertheless have a conjugal life full of meaning, in both human and Christian terms. Their marriage can radiate a fruitfulness of charity, of hospitality, and of sacrifice".

⁷⁰ As Anscombe writes "[f]or if that [reproduction] is not its fundamental purpose [of sex] there is no reason why for example 'marriage' should have to be between people of opposite sexes." See Elizabeth Anscombe, *Contraception and Chastity*, ORTHODOXY TODAY (last visited Dec. 24, 2013), <http://www.orthodoxytoday.org/articles/AnscombeChastity.php> (quoting § 1, para.5).

loving union of partners does not necessarily have to possess or express a procreative function or intention, then arguments could be made that the other characteristic functions/ends of marriage (mutual aid, the responsible orientation of sexual desire) are possible within same-sex relationships and infertile opposite-sex relationships.

Anscombe's prediction about the consequences of the logical separation of the ends of marriage has in some ways been gradually incorporated into the jurisprudence of the U.S. courts in relation to civil marriage, much to the frustration of some traditionalist proponents of opposite-sex only civil marriage. The development of U.S. case law in recent decades has gradually acknowledged the separability of the procreative and companionate ends of civil marriage in relation to heterosexual marriage. The jurisprudence of the U.S. state and federal courts has changed from one that has always principally viewed the institution of civil marriage as a reproductive unit towards an understanding that allows the other purposes of marriage to have a distinct and distinguishable value.⁷¹ Early indications of this direction of travel include *Griswold v. Connecticut* (1965)⁷² which would assert the right to marital privacy over a state interest in barring contraception among married couples.

This was succeeded by other cases that had a more relevant impact on how case law treated non-procreative marital relations, such as *Turner v. Safley* in 1987.⁷³ In this pivotal case the U.S. Supreme Court overturned a Missouri state ban on prisoner inmates marrying, even if consummation (and therefore procreation) was not always possible.⁷⁴ The majority opinion (authored by Justice Sandra Day O'Connor) outlined the significance of marriage in constitutional law as relating not only to procreation, but inferred that independently of this procreative justification civil marriage could be justified by the "emotional support and public commitment ... [as well as the religious or] spiritual significance" of civil marriage to the citizen (in this case the inmate).⁷⁵

Recent court jurisprudence on marriage has tended to avoid the prior U.S. case law precedent of making procreation "the central organizing

Indeed, a prominent Roman Catholic bishop in the UK (Philip Egan of Portsmouth) has noted that the use of contraception and the consequent acceptance of a clear distinction between the unitive and procreative functions of sex within opposite-sex relationships has created, in Egan's words, the "inevitable outcome" of SSCM. See 'Contraceptive Mentality Led to Gay Marriage, Says Egan', THE TABLET, Aug. 3, 2013, at 36.

⁷¹ Consider Charles Reid's analysis. See *supra* note 62 (collecting Reid's articles).

⁷² 381 U.S. 479, 485-486 (1965).

⁷³ 482 U.S. 78 (1987).

⁷⁴ *Turner v. Safley* has relevance in the UK as well. For as Leslie Green has helped clarify, the case law on achieved consummation is ambiguous at best, and the key issue for the law in the UK has been held to be the consenting nature of the union. According to Green, the *capacity* to have sex at all - rather than procreation - is the basis of UK consummation case law. See Green, *supra* note 57, at 14-16.

⁷⁵ 482 U.S. 78, 95-96 (1987).

principle” of marital law.⁷⁶ Legal theorists have used this to build a case for same-sex marriage and the Courts themselves have taken up the development of case law in this area to extend the separability of the goods and purposes of civil marriage to the area of SSCM. This has happened most notably at the state level in *Goodridge v. Dep't of Pub. Health*⁷⁷ and at the federal level in the various opinions handed down in the *Perry* cases.⁷⁸ In these opinions the Courts have not granted the argument that there is a decisive legitimate state interest in ensuring that civil marriage is restricted to heterosexuals on the basis of ‘responsible procreation’ (a variation on O’Brien’s argument), partly because the Courts point to the many infertile or non-procreative heterosexuals who are granted marriage licenses.⁷⁹

This understanding of civil marriage (as involving separable marital goods) from the public political culture is subject to Rawls’s process of political constructivism⁸⁰ using the idea of reflective equilibrium. We may model this in the case at hand by arguing that citizens will reflectively consider, in the original position, what the demands of fairness require in terms of the inclusion of a right to civil marriage within the suite of civil rights included in Rawls’s first principle of justice.⁸¹ The extent and scope of this right is further examined to see whether it applies to opposite-sex

⁷⁶ See Reid, *Marriage in its Procreative Dimension*, *supra* note 62, at 484 (emphasis added).

⁷⁷ 798 N.E.2d 941 (Mass. 2003).

⁷⁸ At the federal level, see *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921 (N.D. Cal., 2010); *aff'd*, *Perry v. Brown*, 671 F.3d 1052, 1063 (9th Cir. 2012) (Reinhardt, J.) (finding California’s Proposition 8, which became CAL. CONST. art. I, § 7.5, unconstitutional under the Equal Protection Clause of the U.S. Constitution’s Fourteenth Amendment), *vacated on other grounds*, *Perry*, 133 S. Ct. 2652 (2013).

⁷⁹ As the *Goodridge* majority found, “[t]he judge in the Superior Court endorsed the first rationale, holding that ‘the state’s interest in regulating marriage is based on the traditional concept that marriage’s primary purpose is procreation.’ This is incorrect. Our laws of civil marriage do not privilege procreative heterosexual intercourse between married people above every other form of adult intimacy and every other means of creating a family. General Laws c. 207 contains no requirement that the applicants for a marriage license attest to their ability or intention to conceive children by coitus. Fertility is not a condition of marriage, nor is it grounds for divorce. People who have never consummated their marriage, and never plan to, may be and stay married. See *Franklin v. Franklin*, 154 Mass. 515, 516 (1891) (‘The consummation of a marriage by coition is not necessary to its validity’). People who cannot stir from their deathbed may marry. See G. L. c. 207, § 28A. While it is certainly true that many, perhaps most, married couples have children together (assisted or unassisted), it is the exclusive and permanent commitment of the marriage partners to one another, not the begetting of children, that is the sine qua non of civil marriage”. *Goodridge*, 798 N.E.2d 941 at 961-62 (references suppressed).

⁸⁰ See RAWLS, *supra* note 8, particularly Lecture III – ‘Political Constructivism’, and Lecture VIII – ‘The Basic Liberties and their Priority’.

⁸¹ Rawls’s first principle of justice states that: “[E]ach person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all”. See RAWLS, *supra* note 8, at 291.

couples only, or is sex-neutral.⁸² Citizens in the original position (which is a heuristic, not a ‘real life’ scenario) do not know their gender, sexual orientation or fertility, as this is cloaked behind a notional veil of ignorance in Rawls’s theory.

One can use the Rawlsian method of a (narrow) reflective equilibrium to illustrate how a typical approach to the civil rights of gay and lesbian citizens might have changed since, say, the 1950’s in the United States. Reflective equilibrium in the political domain is the process by which a person’s initial moral views or intuitions about matters of justice (and/or their application) are tested and adjusted into a coherent equilibrium by comparing them to other relevant background beliefs the person may have, whether these beliefs relate to assumptions (moral or empirical) about human nature or social norms.

In the case at hand, U.S. citizens in the 1950’s might have widely assumed that gay people were rapacious, incapable of fidelity in relationships and suffered from a psychiatric illness. Popular views in the U.S., however, will likely have evolved in recent times to a commonplace understanding that same-sex couples are capable of forming stable and loving relationships in ways that clearly resemble committed opposite-sex relationships.⁸³ Wider changes to sexual mores in relation to the acceptability of (fertile) opposite-sex couples choosing not to procreate are also relevant to this evolved narrow reflective equilibrium in relation to matters of justice and the family.

But would this position also justify going beyond sex-neutral civil marriage towards the more radical option of ‘plural marriages’—perhaps on the putative basis that the distinctive good of mutual aid can be rendered between multiple partners?⁸⁴ I argue that this would not necessarily be the case as the notion of plural marriages and the relations they contain are closer to the generic notion of friendship than the distinctive nature of the *marital* bond, which has been thought of in the western tradition as being dyadic. There are reasons, good reasons in my view, why this dyadic

⁸² To be clear, I do not subscribe to Rawls’s notion of “Justice as Fairness” and the “original position”, which is his own species of political constructivism, though I do take the view that it is helpful in a theory of liberal constitutionalism to have principle(s) that would help citizens reasonably filter out arbitrary or unduly partial motivations, interests or biases.

⁸³ Judge Richard A. Posner argues, without referring to the method of reflective equilibrium, that evolving public views about same-sex relationships, as much as decisions by state and federal courts, have been instrumental to the gradual extension of civil rights to gay and lesbian citizens. See Richard A. Posner, *How Gay Marriage Became Legitimate: A Revisionist History of a Social Revolution*, NEW REPUBLIC (July 24, 2013) <http://www.newrepublic.com/article/113816/how-gay-marriage-became-legitimate>.

⁸⁴ The political theorist Elizabeth Brake is an advocate of this stance – see Elisabeth Brake, *Minimal Marriage*, 120 ETHICS 302 (2010), particularly at sections II and III. Brake believes that permitting plural marriages (or minimal marriages as she calls them) is the only form of civil marriage that is consistent with Rawlsian liberalism properly considered.

understanding of civil marriage is fully defensible in public reasoning and that there is no irresistible logic to the case for plural civil marriage.⁸⁵ ‘Slippery slope’ arguments are often used by those who seem to be on the losing side of a political debate as a tactic to resist change to the *status quo*. They are often not arguments that directly address the issue under scrutiny, but raise fears about undemonstrated further consequences. But all this is not to say that citizens would not, in the original position, consider whether the fulfillment of certain primary goods for citizens might justify some form of legal recognition for certain dependent relationships in particular circumstances. This does not equate to the recognition of plural marriage.

Proponents of SSCM often point to the fact that U.S. citizens have a legal right to civil marriage, currently restricted to opposite-sex couples in most state jurisdictions. There is disagreement about how this situation should be rectified in relation to same-sex couples. Some SSCM advocates argue that the denial of same-sex civil marriage rights is contrary to the Fourteenth Amendment of the U.S. Constitution, with its guarantee that citizens should receive: (1) the Equal Protection of the laws and/or (2) protection under the fundamental rights understood to flow from the Due Process clause of the same amendment. The argument is that there is no clear and rationally compelling state interest that should justify the denial of a civil marriage license to a properly qualifying same-sex couple. This breach of fundamental rights argument can either be approached from the perspective of sex discrimination⁸⁶ or as a specific determination to address unjust discrimination on the basis of sexual orientation.

The same end can be arrived at through the use of Rawls’ notion of the primary good of the ‘social bases of self-respect’—”understood as those aspects of basic institutions normally essential if citizens are to have a lively sense of their worth as persons and to be able to advance their ends with self-confidence.”⁸⁷ Allowing couples currently denied access to the basic legal institution of civil marriage through SSCM would affirm those couples’ sense of self-respect and enable them to live out their lives as citizens more fully. This Rawlsian emphasis on promoting the self-worth of citizens

⁸⁵ For arguments that SSCM does not lead to a slippery slope in recognising plural marriage, see the arguments of Maura I. Strassberg, *Distinctions of Form or Substance: Monogamy, Polygamy and Same-Sex Marriage*, 75 N.C. L. REV. 1501 (1997); James M. Donovan, *Rock-Salting the Slippery Slope: Why Same-Sex Marriage is Not a Commitment to Polygamous Marriage*, 29 N.KY. L. REV. 521 (2002); and Ruth K. Khalsa, *Polygamy as a Red Herring in the Same-Sex Marriage Debate*, 54 DUKE L.J. 1655 (2004-2005). On the more theoretical and historical basis of reasons provided against plural marriages see JOHN WITTE JR., *WHY TWO IN ONE FLESH? THE WESTERN CASE FOR MONOGAMY OVER POLYGAMY* (forthcoming, Oxford University Press).

⁸⁶ Andrew Koppelman, *Defending the Sex Discrimination Argument for Lesbian and Gay Rights: A Reply to Edward Stein*, 49 UCLA L. REV. 519 (2001); Andrew Koppelman, *Why Discrimination against Lesbians and Gay Men Is Sex Discrimination*, 69 N.Y.U. L. REV. 197 (1994).

⁸⁷ See RAWLS, *supra* note 9, at 59.

accords with the way that the U.S. Supreme Court has affirmed that “[f]rom its founding, the Nation’s basic commitment has been to foster the dignity and well-being of all persons within its borders.”⁸⁸ We can see the potential recognition of SSCM as, far from being a “therapeutic” measure to fulfill trivial subjective desires,⁸⁹ as being one that affirms the dignity of its citizens by according them equal access to a basic civic institution.

In what sense does civil marriage generate goods that accrue not only to the spouses but to the wider society, even when procreation is taken out of the picture? For an answer to this question O’Brien could have consulted Judge Walker’s judgment in *Perry v. Schwarzenegger* (which O’Brien cites) that clearly summarizes the government interest in civil marriage apart from social reproduction—including the mix of public and personal benefits that can in some way be attributed to the state of civil marriage.⁹⁰ Benefits which generate political value include allowing the civil law to effectively organize matters such as hospital visiting rights, tenancy rights on the death of one partner, inheritance entitlements and so on. All of which have the benefit of avoiding unnecessary legal disputes or queries in cases when one spouse dies or is seriously ill. Without these clearly established rights the prospect of dispute and the attendant costs to public agencies (and family members) is considerably more likely. This is surely of some social and political value.

We may also say that marital associations that encourage supportive relationships and cohabitation prevent the proliferation of single person households with the negative externalities generated by them. Inversely, these positive social externalities from civil marriage insofar as it helps sustain relationships may broadly be classified as those that relate to: 1) the public health improvements that may accrue from the maintenance of stable relationships;⁹¹ 2) the lower social care costs that fall on the state as a result of cohabiting partners assisting each other; and 3) the lower level of pressure on scarce housing stock that fewer single person households would ensure.⁹²

More broadly still, the virtues facilitated by marital bonds with their characteristic habits of commitment and fidelity can bleed into political virtues that create political value. Marriage can be a ‘school of virtue’ in

⁸⁸ *Goldberg v. Kelly*, 397 U.S. 254, 264–65 (1970), which was a case relating to the provision of welfare.

⁸⁹ O’Brien writes of the recognition of SSCM as an example of the “therapeutic” role of government which he feels is illegitimate. *See supra* note 1, at 436.

⁹⁰ *Perry*, 704 F. Supp. 2d, at 961–63.

⁹¹ *Id.* at 962 (referencing public health evidence on this claim). Healthier people in a society clearly generate political value in terms of greater economic and social productivity and the benefit of lower health care costs, some of which would certainly fall on the taxpayer.

⁹² These costs are widely known, but a summary of them can be found at: JIM BENNETT & MIKE DIXON, SINGLE PERSON’S HOUSEHOLDS AND SOCIAL POLICY: LOOKING FORWARD (2006), available at <http://www.jrf.org.uk/sites/files/jrf/bennett-9781859354759.pdf>.

which spouses learn patience, listening skills, forbearance, compromise, mutual understanding, commitment and mutual care ‘in sickness and in health’. These personal virtues overlap with important political virtues for citizens, as citizens are called to embody many of the same virtues in a different mode as ‘civic friends’. These virtues or capacities are not strictly limited to the twin Rawlsian capacities to form a plan of life (the ‘non-public’ rational power) and a sense of political justice (‘public’ reasonableness). O’Brien should not be surprised at such arguments as they have been part of the public political culture of western societies for centuries. The legal historian and jurist John Witte, for example, refers to the:

core insight of the Western tradition - that marriage is good not only for the couple and their children, but also [good] for the broader civic communities of which they are a part. The ancient Greek philosophers and Roman Stoics called marriage “the foundation of the republic,” “the private font of public virtue.” The Church Fathers called marital and familial love “the seedbed of the city,” “the force that welds society together.” Catholics called the family ... “a kind of school of deeper humanity.” Protestants called the household ... a “little commonwealth.” American jurists ... taught that marriage is both private and public, individual and social ..., a useful if not an essential association, a pillar if not the foundation of civil society. At the core of all these metaphors is a perennial Western ideal that stable marriages and families are essential to the....flourishing, and happiness of the greater commonwealths...⁹³

Moreover, as the natural lawyer Gary Chartier has argued, one does not have to see the relationships included in marital law as being universally morally admirable to recognize that societal benefits may accrue from their legal recognition.⁹⁴

Towards the end of his article,⁹⁵ O’Brien reaches for the idea of legally recognized domestic dependency arrangements, almost as a *deus ex machina*, claiming that this addresses the interests of mutually reliant same-sex couples or other caring relationships not based on gender or orientation (e.g. two siblings living together). He does not spell out how these domestic dependency rights would differ from the rights accorded to civilly married couples (power of attorney, tax benefits, hospital visitation rights etc.) This move from O’Brien is a significant concession, because in doing so he acknowledges that mutual aid and companionate aspects of a relationship

⁹³ See Witte, *supra* note 58, at 1070.

⁹⁴ As Chartier writes, those “who believe same-sex relationships are wrong must recognize that such relationships exist and will continue to exist for the foreseeable future. Societal pressure may drive these relationships underground, but it is unlikely to eradicate them. Given that they exist, societies, even those that regard them with distaste, cannot simply ignore them. Providing people involved in such relationships with the option of marriage will help them to contribute to each other's welfare and to that of society as a whole.” See Gary Chartier, *Natural Law, Same-Sex Marriage, and the Politics of Virtue*, 48 UCLA L. REV. 1622 (2000-2001).

⁹⁵ O’Brien, *supra* note 1, at 452-53.

are clearly a political value—otherwise there would be no reason to allow for domestic dependency arrangements as he proposes.

Various inconsistencies arise with this concession. Applying Ockham's razor to the logic of O'Brien's argument regarding the imperative of social reproduction, taken as the sole or dominant political value, would surely be to abolish civil marriage altogether (as there is no stipulation or specific incentive to procreate *inherent* in traditional marriage law) and replace it outright with universal domestic dependency arrangements (open to all different types of couples, sexual relationships or not) accompanied by further tax or other financial incentives for those couples in domestic dependency arrangements who *actually go on to have children*. Monetary benefits to such couples could vary depending on the demographic challenges facing a particular society. A society with a fertility rate above replacement level might not need to put in place any such incentives and could use the resources saved for other socially useful purposes. Those societies below replacement level could make a choice as to whether immigration or specific tax incentives for citizens to have children would be the most socially or politically appropriate response to the challenges of an unduly ageing population.

This solution, from the logic of O'Brien's position (which is not my own), would also mean that the imperative of treating 'like cases alike' would be met—infertile or non-procreative hetero- and homosexual couples would have the same legal status (domestic dependency status) while couples who *actually* contributed to the reproduction of a society by having their own children would have additional recognition from the state (domestic dependency status plus financial benefits for child rearing).⁹⁶ Yet O'Brien does not reach for this obvious solution despite his professed theoretical commitment to moral neutrality.⁹⁷

Equally, O'Brien's domestic dependency proposal raises the following question: if virtually all the legal rights and privileges of marriage are granted to same-sex couples through this domestic dependency route, is a society granting same-sex couples (and others) the *de facto* equivalent of civil marriage? In which case if the *only* thing that he is seeking to withhold from same-sex unions is the 'expressive' approval of the law contained in social meanings of marriage—something at stake in the *Perry* case—then this seems quite a semantic approach to the question of SSCM. If the basic legal entitlements of civil marriage are to be conferred to same-sex couples via the domestic dependency route, why not go the whole way, especially given the majority popular support that has emerged for SSCM

⁹⁶ It is worth pointing out that a number of European nations have specific welfare benefits relating to the number of dependent children within a family (including single parent families). In the United Kingdom this benefit is known as Child Benefit.

⁹⁷ Whether this commitment to moral neutrality is O'Brien's own personal view or a 'stage piece' is not clear, as I shall explain in §V of this article.

in the United States in recent times?⁹⁸ The fact that a number of religious organizations in the U.S. consider such couples to be civilly married (not just ‘domestic dependents’) and wish to be free to bless or solemnize those relationships must surely add a further reason to consider going the whole way to recognizing SSCM.⁹⁹

IV. SAME-SEX MARRIAGE, ‘RATIONAL BASIS’ REVIEW AND U.S. CONSTITUTIONAL LAW

Matthew O’Brien raises the issue of the appropriateness of the federal court’s use of the rational basis review test to strike down laws that bar same-sex couples from access to civil marriage (in the *Perry* case). He sees this as a doubly mistaken move in that not only have the Courts: (1) made the wrong substantive decision but, in doing so, they have (2) directly rejected what he considers to be the *only* rational governmental interest in civil marriage *per se* (i.e. societal reproduction).

I hold that O’Brien’s complaint on juristic grounds in this respect is overblown, even setting aside the substantive theoretical issues already discussed. I say this firstly on the basis that the Ninth Circuit Court of Appeals and the U.S. Supreme Court’s rulings in the *Perry* case (and the *Goodridge* decision in Massachusetts for that matter) have actually not been nearly as expansive and wide ranging as they might have been, in that neither decision announced a nation-wide constitutional right to SSCM by using the fundamental rights doctrine derived from the Due Process Clause of the Fourteenth Amendment.

Second, I would question the parallel O’Brien draws between the notion of a rational government interest (held by the states), which is crucial in determining the outcome of a rational basis review by the Federal Courts, with aspects of Rawls’s political liberalism. Taking the latter point first, at certain points in his article O’Brien seeks to deliberately parallel the rational basis test in U.S. law with John Rawls’s notion of public reason.¹⁰⁰ Though there are superficial similarities between the two ideas the connection between them may not be useful in the analysis of the same-sex marriage issue.

There are again two reasons why the parallel is not particularly apt. First, the rational basis test is a specific *legal* principle/precedent from U.S. constitutional law grounded in the historical experience of that polity relating to the respective roles of the legislative and judicial branches and the

⁹⁸ See *Gay Marriage: Key Data Points from Pew Research*, PEW RESEARCH CENTER (June 11, 2013), <http://www.pewresearch.org/2013/03/21/gay-marriage-key-data-points-from-pew-research/> (showing that a plurality of people surveyed since 2011 by the Pew Center in the U.S. favor SSCM).

⁹⁹ As Leslie Green reminds us, Canada took the step to legalize SSCM when an Ontario court retroactively recognized two religious same-sex marriages as legally valid. See Green, *supra* note 57, at 7-8.

¹⁰⁰ O’Brien, *supra* note 1, at 413, 416, 463.

structuring of powers between the federal and state governments (where the latter have general ‘police powers’, as O’Brien notes).¹⁰¹ Judges Walker and Reinhardt do not themselves draw any implicit or explicit parallel between their rulings on Proposition 8 and Rawls’s theory, despite O’Brien implication that there is a link between them.¹⁰²

Secondly, nowhere in *Political Liberalism* or *Justice as Fairness: A Restatement*, as far as I can determine, did Rawls himself draw a parallel between the rational basis review and his concept of public reason.¹⁰³ Rawls’s reference to the U.S. Supreme Court as normatively being the epitome of public reason in *Political Liberalism*¹⁰⁴ does not imply that Rawls saw public reason and rational basis review as analogous. O’Brien seems to allow the parallel between the rational basis review and notion of public reason do work in his article that, on closer consideration, it cannot.

The history of the *Perry* case may also not be helpful in clarifying the key issues in relation to SSCM and Rawlsian public reason. This is for two reasons: first because the findings of fact from Judge Walker showed that the *actual* reasoning put forward by the Proposition 8 proponents in court were not arguments closely tethered to legitimate ‘rational’ government interests as he judged them, but were arguments that were either irrelevant (such as those addressing the desirability of gay parenting – parenting rights were not affected by the passage of Proposition 8), arbitrary (such as the argument that infertile heterosexual couples should be able to marry but same-sex couples must not) or those grounded in *animus* against same-sex couples¹⁰⁵ (the ‘findings of fact’ documented the formal participation in the California ‘Project Marriage’ campaign of persons publicly arguing that

¹⁰¹ These general powers reside in the States rather than the Federal Government. Cf. U.S. CONST. art. IV, § 4, which guarantees a “Republican Form of Government” to the States.

¹⁰² O’Brien, *supra* note 1, at 413.

¹⁰³ We may reasonably assume that that John Rawls would have been aware of the notion of ‘rational basis review’ in U.S. law, which was widely known as a key judicial instrument in racial and gender cases for much of the twentieth century. *But cf. infra* note 107.

¹⁰⁴ See RAWLS, *supra* note 8, at 231.

¹⁰⁵ Like the jurist Michael J. Perry (a Catholic proponent of SSCM), I do not assume that those opposed to same-sex marriage are generally motivated by animus or bigotry (though some SSCM opponents certainly appear to hold bigoted views). I agree with Perry in regretting that the federal courts, including the US Supreme Court in its ruling on the *Windsor* case, has seemingly deemed that opposition to SSCM is fundamentally premised on *animus* rather than any (non-bigoted) ethical argument, however morally mistaken and/or politically unjust such arguments may be. See Michael Perry, *Right Decision, Wrong Reason: Same-Sex Marriage & the Supreme Court*, COMMONWEAL (Aug. 5, 2013), <http://www.commonwealmagazine.org/right-decision-wrong-reason>. This does not mean that I echo the (typically) splenic rhetoric of Justice Antonin Scalia who, in his dissenting opinion in the United States v. Windsor, hyperbolically accuses the majority of treating SSCM opponents as ‘enemies of the human race’. See *Windsor*, 133 S. Ct. 2675, 2709 (2013) (Scalia, J., dissenting).

homosexuals were prevalently child sex abusers who would lead America into the hands of the Evil One if allowed to marry).¹⁰⁶

It is of course possible to hold that there are clear public reasons (in Rawlsian terms) for the recognition of SSCM, whilst not believing that any particular legislative decision in relation to SSCM has failed a rational basis review in U.S. constitutional practice.¹⁰⁷ Moreover, one can hold the general normative view—as I do—that that decisions relating to matters such as SSCM should generally not be made by the judicial branch but should be the province of legislatures (a view often described as ‘popular’ or ‘political’ constitutionalism).¹⁰⁸ In other words, even if O’Brien succeeds in persuading his critics that the rational basis test was wrongly applied in this or that case in the federal courts, he would still not necessarily have succeeded in his aim of demonstrating that there are no Rawlsian public reasons in favor of the recognition of SSCM.¹⁰⁹

In any case O’Brien’s fears about the application of the rational basis test to SSCM cases are exaggerated because the basis of the Ninth Circuit Appeals Court’s decision in favor of the plaintiffs in the *Perry* case was limited and *sui generis*. The Ninth Circuit decided not that the constitution grants a general civil right to SSCM (as *Loving v. Virginia*¹¹⁰ established a civil right under the Fourteenth Amendment to race-blind marriage) but ruled that the *withdrawal* of SSCM brought in by Proposition 8 was not—

¹⁰⁶ “[T]he America Return to God Prayer Movement, which operates the website ‘1man1woman.net.’ ... 1man1woman.net encouraged voters to support Proposition 8 on grounds that homosexuals are twelve times more likely to molest children [according to 1man1woman.net] ... and because Proposition 8 will cause states one-by-one to fall into Satan’s hands”, *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921, 937, 989 (N.D. Cal. 2010), at 22.

¹⁰⁷ Of relevance here is the view of U.S. Attorney General Holder that “classifications based on sexual orientation should be subject to a heightened standard of constitutional scrutiny under equal protection principles” as are classifications such as race, national origin, alienage, non-marital parentage, and gender (in varying degrees) than the rational basis test. Letter from U.S. Attorney General, Eric H. Holder, Jr., to Speaker of the U.S. House of Representatives, John A. Boehner, regarding *McLaughlin v. Panetta*, Civ. A. No. 11-11905 (D. Mass., Feb. 17, 2012). In this view, the SSCM issue should be treated under a different rubric than the straightforward rational basis test, that of ‘heightened scrutiny’.

¹⁰⁸ For an example of such a view see RICHARD BELLAMY, *POLITICAL CONSTITUTIONALISM* (2009) or, from a different perspective, the work of Mark Tushnet including in MARK TUSHNET, *TAKING THE CONSTITUTION AWAY FROM THE COURTS* (2000).

¹⁰⁹ As Robin West writes “[t]his decision is thus of no relevance to cases challenging a state’s refusal to extend marriage to include gays and lesbians [generally], and it is of no relevance to cases challenging a state’s withdrawal of such a right if that right is also accompanied by a denial of concrete benefits and accompanied by some explanation—such as the superiority of heterosexual parenting—for the decision to do so. *Perry v. Brown* is nothing more than a *sui generis* decision for a unique set of facts.” Robin L. West, *A Marriage Is a Marriage Is a Marriage: The Limits of Perry v. Brown*, 125 HARV. L. REV. F. 48 (2012).

¹¹⁰ 388 U.S. 1 (1967).

in the circumstances of the specific campaign and legal context of California—clearly advanced by its proponents to serve a defined and legitimate government interest. It was thus considered a breach of equal protection of the laws under the Fourteenth Amendment by Judge Reinhardt. In limiting the judgment to the Equal Protection Clause, it overruled Judge Walker’s District Court ruling that Proposition 8 contravened *both* the fundamental rights jurisprudence (from the Due Process Clause) as well as the Equal Protection Clause.

As such the controlling precedent for the case for the Ninth Circuit was not the chain of case law emerging from the U.S. Supreme Court’s establishment and elaboration of a constitutional right to civil marriage (in *Loving*, *Zablocki v. Redhail*,¹¹¹ *Turner v. Safley*, etc.) but the U.S. Supreme Court’s much more limited judgment in *Romer v. Evans* to strike down a Colorado law that withdraw protection from discrimination against citizens purely on basis of their sexual orientation.¹¹² A decision directly based on *Loving*, *Zablocki* and *Turner* using the fundamental rights doctrine based on the Fourteenth Amendment might have led to a move to enforce a sex-neutral legal definition of civil marriage across the states. In the event, the *Perry* case was vacated as the U.S. Supreme Court decided, by a majority decision, that the plaintiffs did not have legal standing in the case (in part because the State of California did not contest the lower court’s rulings). The decision of the lower courts striking down Proposition 8 thus stood.¹¹³

The majority decision of the U.S. Supreme Court in the twined case of *United States v. Windsor*¹¹⁴ also did not affirm a constitutional right to sex-neutral civil marriage. *Windsor* was, again, decided on narrow grounds. First, on the basis that it is state governments and not Congress who have traditionally decided matters relating to marital and family law and, second, on the Fifth Amendment’s Due Process requirement.¹¹⁵ (Edith Windsor’s lawyers did not invoke the fundamental rights jurisprudence of the Fourteenth Amendment to support her claim). As Chief Justice John Roberts noted in his dissent, the majority decision in *Windsor* leaves intact Section II of Defense of Marriage Act, which legislates that states that do

¹¹¹ *Zablocki v. Redhail*, 434 U.S. 374 (1978).

¹¹² *Romer v. Evans*, 517 U.S. 620 (1996). This case, in effect, introduced a rational basis test ‘with bite’, which is perhaps something of a half-way house between a conventional rational basis test and a heightened scrutiny review. This ‘with bite’ element of the rational basis test is not mentioned by O’Brien but is perhaps relevant in considering the *Perry* case. Cf. Jennifer Serrine, *A Rational Approach to California’s Proposition 8*, 9 DARTMOUTH L.J. 49 (2011).

¹¹³ *Perry*, 133 S. Ct. 2652 (2013).

¹¹⁴ 133 S. Ct. 2675 (2013) (striking down Section III of the Defense of Marriage Act, 1 U.S.C. § 7 (1996), which barred the federal government and its agencies from recognizing the civil marriages of same-sex couples from states where SSCM is legal).

¹¹⁵ Which mandates, according to the U.S. Supreme Court’s prior jurisprudence, the equal protection of the law for citizens under federal authority. See *Bolling v. Sharpe*, 347 U.S. 497, 499 (1954).

not recognize SSCM are not compelled to recognize the civil marriages of same-sex spouses from states that do.¹¹⁶ The U.S. Supreme Court thus again refrained from using the chain of precedent from *Loving*, as the Ninth Circuit did with regard to *Perry*.

The decision of the U.S. Supreme Court to stay the December 20, 2013 ruling of the U.S. District Court for the District of Utah striking down the ban on SSCM contained in that state's constitution (on the basis of both the Fourteenth Amendment's fundamental rights doctrine and the Equal Protection Clause) may indicate that the highest court is not yet of the view that the denial of SSCM by a state is a contravention of the federal constitutional right to marry (*à la Loving* and marital race bars).¹¹⁷

The question of what can serve as a legitimate government interest goes to the heart of the traditional powers to enforce public morality under the general 'police powers' of state governments. This is clearly a controversial question on which legal and social attitudes have changed over the decades (for instance in relation to the legal censorship of the content of theatrical performances or movies). Key to the discussion of the enforcement of public morality is surely its interplay with the contrasting doctrine of the right to privacy that the U.S. Supreme Court has found, in the post-war period, in the 'penumbras' of the Bill of Rights.¹¹⁸ There are, of course, other ways than a general and wide-ranging right to privacy to conceive of the limits to the legitimate scope of government in advancing human goods (or militating against human wrongs)¹¹⁹ or in justifying the existence of

¹¹⁶ 133 S. Ct. at 2696 (Roberts, C.J., dissenting) ("The Court does not have before it, and the logic of its opinion does not decide, the distinct question whether the States, in the exercise of their 'historic and essential authority to define the marital relation,'... may continue to utilize the traditional [heterosexual] definition of marriage." (quoting the opinion of the Court)).

¹¹⁷ Order in pending case, *Herbert v. Kitchen*, Docket No. 13A687, Jan. 6, 2014, 571 U.S. ___. The Tenth Circuit Court of Appeals is seeking briefing on the case as I write. A similar case was decided on January 14, 2014 by the U.S. District Court for the Northern District of Oklahoma, striking down that state's constitutional ban on SSCM and disapplying Section II of the Defense of Marriage Act (*Bishop v. United States*, 4:04-cv-00848-TCK-TLW (N.D. Okla.)). The U.S. District Court for the Northern District of Oklahoma has stayed its ruling pending the Tenth Circuit's ruling regarding the *Kitchen v. Herbert* case.

¹¹⁸ Most famously in *Griswold v. Connecticut*, 381 U.S. 479, 484 (1965), Justice Douglas, delivering the opinion of the Court, stated that there is a right of marital privacy that justifies the legalisation of contraceptives for married couples on the basis that "specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance. ... Various guarantees create zones of privacy." Douglas based these (unenumerated) guarantees in the *Griswold* case on the penumbras of the First, Third, Fourth, Fifth and Ninth Amendments.

¹¹⁹ A Millian may wish to refer to 'the harm principle' and contemporary Thomists use Aquinas's distinction between sin and crime and his recommendation that it is necessary only to legally enforce matters of morality that relate to interpersonal justice, civil peace or public order. See for instance the limited role of government outlined in John Finnis, *Public Good: The Specifically Political Common Good in Aquinas*, in *NATURAL LAW AND MORAL INQUIRY: ETHICS, METAPHYSICS, AND POLITICS IN THE WORK OF GERMAIN*

‘liberty’ or ‘privilege’ rights grounded in the autonomy of the legal person,¹²⁰ however autonomy is conceived from the philosophical perspective.

Concerns expressed by O’Brien that the U.S. Supreme Court has endorsed a comprehensive ethical conception of autonomy in its recent case law may also be somewhat exaggerated¹²¹—though I would concur that, when taken in isolation, some of the purple passages authored by Justices O’Connor and Kennedy in the *Planned Parenthood* and *Lawrence* majority opinions respectively¹²² are regrettably apt for such an interpretation. It is

GRISEZ, (Robert P. George ed., 1998) and John Finnis, *Liberalism and Natural Law Theory*, 45 MERCER L. REV. 687 (1994).

¹²⁰ The intellectual historian Brian Tierney argues that as early as the twelfth century canon lawyers were referring to *ius* (right) as a personal faculty (*facultas*) or moral power (*potestas*). The relevant aspect of Tierney’s work in this context is the argument that the concept of explicitly subjective ‘liberty’ (‘privilege’) or ‘power’ rights (in the Hohfeldian sense) extended back to the 1180’s where English canonists were referring to a concept of a ‘permissive natural law’ based on a subject’s faculty for choice in situations where actions “are neither commanded nor forbidden by the Lord or by any Statute” (quoting the English canonist author of *Summa, In Nomine*). See BRIAN TIERNEY, THE IDEA OF NATURAL RIGHTS 67 (2001). Also relevant here is: Jean Porter, *Natural Right, Authority and Power, the Theological Trajectory of Human Rights*, 3 J.L. Phil. & Culture 299 (2009). I am not, of course, suggesting here that either Tierney or Porter thinks that ‘unenumerated’ substantive due process rights, as found in twentieth century U.S. Supreme Court jurisprudence, are present in English medieval canonistic thought, only that such twelfth and thirteenth century jurists and philosophers had already begun to see that personal self-dominium (practical self-direction if you will) inferred a zone of liberty that should be protected. This is perhaps an early conceptual root of the modern right of privacy, even if we may disagree on how such privacy rights should be specified in relation to controversial moral issues such as abortion or sexual relations.

¹²¹ Relevant here in interpreting the notion of autonomy found in the *Planned Parenthood* and *Lawrence* opinions is James E. Fleming’s notion of “deliberative autonomy” which he sees as much narrower than a comprehensive (in the Rawlsian sense) liberal or libertarian understanding of autonomy. Fleming sees deliberative autonomy as being drawn from the underlying values of freedom of conscience and the other civil liberties contained in the structure and text of the U.S. Constitution. Fleming interprets the *Lawrence* majority opinion in this more limited light; pointing out that it protected same-sex intimate relations on the basis that such conduct can be seen as “closely analogous to heterosexual intimate conduct, which is already constitutionally protected”, see JAMES E. FLEMING, SECURING CONSTITUTIONAL DEMOCRACY: THE CASE OF AUTONOMY 137 (2006). Chapters 5 and 6 of the book give Fleming’s considered view of deliberative autonomy, and the whole work is an attempt to show how Rawlsian political liberalism can be used to create a coherent philosophic reading of the constitution. I cite Fleming’s work not because I share its fundamental jurisprudential position, but because it propounds a well thought through reading of autonomy rights within the U.S. constitutional order—a view that O’Brien does not consider.

¹²² The majority opinion in *Planned Parenthood of Southeastern Pa. v. Casey*, 505 U.S. 833, 851 (1992), stated that “[t]hese matters, involving the most intimate and personal choices a person may make in a lifetime, choices central to personal dignity and autonomy, are central to the liberty protected by the Fourteenth Amendment. At the heart of liberty is the right to define one’s own concept of existence, of meaning, of the universe, and of the mystery of human life. Beliefs about these matters could not define the attrib-

important to point out that neither of these purple passages directly imply that moral norms or the moral law are themselves somehow *created or wholly specified by the citizen in her/his exercise of autonomy*. It is therefore questionable that these opinions have reified in U.S. constitutional law a substantive and comprehensive doctrine of *constitutive* moral autonomy, as O'Brien implies. Moreover, a modest and limited notion of moral autonomy or 'practical freedom' is consistent with a range of reasonable comprehensive ethical, philosophical or religious doctrines, even those that allow a strong role for revealed divine law (such as Catholicism).¹²³ What is important—and contested—is the extent to which the recognition of the autonomy or practical freedom of legal persons mandates a zone of privacy that the law must respect. This zone of privacy itself helps to delineate the permissible scope of 'public morals' legislation.¹²⁴

V. CONCLUSION

In concluding it is worth examining the overarching strong and weak points that come over in O'Brien's provocative article. His article valuably focuses on Rawls's argument that the family has a clear social function in the reproduction of a political society over time. This point has the potential to contribute to public reasoning in favor of the inclusion of a right to civil marriage in the list of civil rights within a liberal political conception of justice. This point has not had the prominence in the literature on same-sex marriage that it should have. Rawls's insights on social reproduction have some relevance, particularly in societies where the birth rate might dip well below replacement level.

utes of personhood were they formed under compulsion of the State." In *Lawrence v. Texas*, 539 U.S. 558, 562 (2003), the purple passage announces: "[f]reedom extends beyond spatial bounds. Liberty presumes an autonomy of self that includes freedom of thought, belief, expression, and certain intimate conduct. The instant case involves liberty of the person both in its spatial and in its more transcendent dimensions."

¹²³ For an instructive survey of how both pre-modern and modern ethical worldviews have posited or valued forms of autonomy, practical freedom or 'self-dominium' see T.H. Irwin, *Continuity in the History of Autonomy*, 54 *INQUIRY: AN INTERDISCIPLINARY JOURNAL OF PHILOSOPHY* 442 (2011) and Brian Tierney, *Dominium of Self and Natural Rights before Locke and After*, in *TRANSFORMATIONS IN MEDIEVAL AND EARLY MODERN RIGHTS DISCOURSE* 176-94 (V. Makinen & P. Korkmann, eds., 2006).

¹²⁴ Even though O'Brien points to what he sees as a fundamental inconsistency between the privacy rights implied in the *Planned Parenthood* and *Lawrence* cases with the notion of public morals legislation (O'Brien, *supra* note 1, at 412-13), the courts themselves have seen no fundamental clash and continue to implement public morals legislation – albeit with a different scope than they would have before the jurisprudence of the U.S. Supreme Court in the post-war period carved out certain privacy rights. O'Brien notes this himself (*id.* at 413) when he quotes Judge Reinhardt's Ninth Circuit opinion in *Perry v. Brown* that affirms the *continuation* of the states' police powers on public morality. (I make no judgment here about the rights and wrongs of the *Planned Parenthood* decision, which relates to abortion.)

O'Brien, however, spoils his argument when he relies on a tendentious reading of Rawls's later oeuvre, claiming that Rawls propounds a thoroughgoing moral neutralism which, in a balanced reading, he does not. This reading of Rawls is used by O'Brien to invalidate arguments for SSCM on the basis that its proponents hold substantively liberal ethical commitments. As I have demonstrated, Rawls did not believe that holding reasonable comprehensive commitments and advancing them in the political (or academic) domain discredits or disqualifies the otherwise publicly accessible arguments that the same people may also advance in favor of SSCM, or any other issue relating to the basic structure of a constitutional democracy.

The particular account of the justification of civil marriage O'Brien proffers does not escape some internal contradictions that make it vulnerable to the charge that it is unduly inconsistent in its treatment of same-sex and opposite-sex couples. His sole justification of civil marriage is that it engenders natural social reproduction and preferable parenting conditions.¹²⁵ Yet O'Brien considers that infertile heterosexual couples, or those with no intention to procreate, should be allowed to marry. If the justification for the civil recognition of marriage is *solely* the procreative function of such a relationship (rather than the broader recognition of loving unions and the polyfunctional ends that are produced by that union), then there is little justification for allowing post-menopausal women to marry or those who have known medical conditions that prevent procreation. O'Brien undermines his own train of argument by affirming that civil marriage should be limited to opposite-sex couples regardless of their *actual* fertility because *in abstracto* they are "characteristically" reproductive and that heterosexual sex is in a generic sense "intrinsically generative".¹²⁶

Despite O'Brien's protestations—which are not argued through dialectically against potential objections—such a position presupposes some form of (contestable) natural philosophy or metaphysical anthropology. Such a metaphysical anthropology would struggle to be expressed in the purely political terms that O'Brien himself interprets Rawls as stipulating within his understanding of public reason. For it is surely far from a statement of the obvious (as O'Brien infers) that a postmenopausal woman wishing to get married is entering an "intrinsically generative" union with her prospective husband. It is biologically 'natural', after all, that people of both sexes (e.g. in childhood and extreme old age in men) to have periods when they are not 'generative'. As is obvious, in women the postmenopausal infertile period is a result of a perfectly normal biological process, it is

¹²⁵ As O'Brien and I are not trained in the methods of child psychology or its allied disciplines (and are writing in a law journal), I shall pass over his strong and substantive judgments about the right conditions for child rearing while noting – as O'Brien does with more than a little chagrin – that the main professional association in the United States in that discipline has formally stated that that children are not generally disadvantaged by being brought up by same-sex parents. See O'Brien, *supra* note 1, at 443.

¹²⁶ O'Brien, *supra* note 1, at 438-39.

not considered either a pathology or a disability.¹²⁷ This is a key flaw in O'Brien's argument that does not receive sufficient argumentative attention given the stakes involved: the failure of the "intrinsically generative" argument would undermine his support for traditionalist heterosexual marriage, or at least mean that age restrictions (or a fertility test) should be considered for opposite-sex partners wishing to get married in civil law.¹²⁸

This lack of argumentation is surprising because reference to "characteristically" reproductive relationships seems to be a variation on the view, put forward by Robert George and John Finnis, that infertile heterosexual sex is still of a reproductive *type* or *kind* and thus infertile heterosexuals should still be eligible to marry. But this parallel is awkward for O'Brien because he has clearly distanced himself from the New Natural Law understanding of ethics and has already judged that their arguments are inapt with regard to public reason.¹²⁹ In this sense the "intrinsically generative" argument for traditionalist marriage advanced by O'Brien is vulnerable to the same criticisms to those made against New Natural Law theorists who claim that *only* married heterosexual (and non-contracepted) coital sex acts can properly be considered morally licit sexual acts.¹³⁰ In my judgment and that of others, criticisms of the New Natural Law argument in relation to the so-called 'sterility objection' have never been answered adequately, despite the spilling of much ink on the question.¹³¹

¹²⁷ It would be more defensible to say that women, in particular, are intrinsically generative at points *in their fertile period* between puberty and menopause. But this would undercut the logic of O'Brien's argument, which (silently) seems to rely on a naturalistic notions of 'natural kinds' or 'finalities' that smack of a 'comprehensive' philosophic conception of gender and sexual potency. Though I am far from unsympathetic with a revised and critically nuanced form of Aristotelian(-Thomistic) ethical naturalism, which seems to be back in academic fashion, I do not draw the same sort of applied moral conclusions from a (neo) naturalistic philosophical anthropology that O'Brien does.

¹²⁸ Martha Nussbaum argues that the procreative argument for traditional heterosexual marriage is inconsistent and challenges its proponents to answer the *reductio* of why there should not be an upper age bar for women wishing to apply for a marriage license to cover their post-menopausal period. See Martha C. Nussbaum, *Reply - A Right to Marry?*, 98 CAL. L. REV. 744 (2010). Needless to say I do not favor such an age bar.

¹²⁹ E.g., O'Brien, *supra* note 1, at 418ff.

¹³⁰ New Natural Law theorists consider all other types of sex acts, including all heterosexual marital but non-coital sex (such as oral sex), to be intrinsically immoral and self-alienating, a position that has come under intense scrutiny and criticism in a number of places, not least in NICHOLAS BAMFORTH & DAVID A. J. RICHARDS, *PATRIARCHAL RELIGION, SEXUALITY, AND GENDER: A CRITIQUE OF NEW NATURAL LAW* (2008).

¹³¹ See on the one side of the argument Girgis et al. *supra* note 53 (who claim that infertile married couples are still 'biologically' ordered to procreation) and on the other Andrew Koppelman, *Careful with That Gun: Lee, George, Wax, and Geach on Gay Rights and Same-Sex Marriage* (Northwestern Public Law Research Paper No. 10-06, Jan. 11, 2010), <http://dx.doi.org/10.2139/ssrn.1544478>, and Erik A. Anderson, *A Defense of the 'Sterility Objection' to the New Natural Lawyers' Argument Against Same-Sex Marriage*, 16 ETHICAL THEORY & MORAL PRAC., 759 (2013).

In my alternative politically liberal understanding of the SSCM issue we do not bypass the widely received western understanding of marriage in civil and canon law as a society founded on interpersonal love for the advancement of the goods of procreation, mutual aid and sexual continence. Following the post-war introduction of contraception we can understand better that the traditional goods of marriage of mutual aid and continence are separable from the reproductive purpose of marriage. The separability of the traditional goods of marriage has also been recognized, by inference, in the public political culture of the U.S. in constitutional case law (such as in *Turner v. Safley*). As I have noted, this changing legal precedent has been accompanied by the rightful marginalization of the hitherto commonly held assumption that gay people are generally rapacious or suffer from a psychiatric illness.

Thus we can, by using a process of Rawlsian political constructivism, come to a liberal political conception of justice that includes same-sex couples as well as opposite-sex couples in the general legal right to civil marriage. There is a substantial gap between arguments expressed in publicly reasonable terms in favor of civil marriage that *includes* the key function of societal reproduction over time and O'Brien's position, which is that the *only* publicly reasonable argument for civil marriage is social reproduction. I find the *inclusive* Rawlsian argument not only publicly reasonable but also persuasive, whereas the latter *exclusive* argument (as advanced by O'Brien) appears to be arbitrary and tinged with ideological undertones.¹³² After all, does O'Brien believe that extending marital rights to same-sex partners would adversely affect the birth rate among married heterosexual couples?¹³³

How might we interpret the impetus behind O'Brien's article? The author is curiously circumspect in that he gives the reader no direct indication as to whether he actually subscribes to Rawls's understanding of political liberalism. This leads one to at least raise the question of whether he is us-

¹³² At times O'Brien's article does not help to dispel the view that he may be writing from a strongly ideological perspective when he draws some colorful (but unfortunate) parallels or *reductions* between arguments to recognize same-sex relationships with, for instance, the mutually dependent relationship between a "pimp and prostitute" or when he claims that "homosexual orientation is on a political par with, say, a traditional order of chivalry". These parallels come in O'Brien, *supra* note 1, at 462 and 459 respectively.

¹³³ Such an implicit argument was rightly dismissed by the United States Court of Appeals for the Ninth Circuit in relation to Proposition 8 when the majority opinion stated that "withdrawing from same-sex couples access to the designation of 'marriage'—without in any way altering the substantive laws concerning their rights regarding childrearing or family formation—will encourage heterosexual couples to enter into matrimony, or will strengthen their matrimonial bonds, we believe that the People of California 'could not reasonably' have 'conceived' such an argument 'to be true.' It is implausible to think that denying two men or two women the right to call themselves married could somehow bolster the stability of families headed by one man and one woman. While deferential, the rational-basis standard 'is not a toothless one.'" See *Perry v. Brown*, 671 F.3d 1088 (9th Cir. 2012) (internal citations omitted).

ing an interpretation of Rawls to pick holes in the case for SSCM, by advancing a *faux* or ironical internal critique of the widely made Rawlsian case for SSCM.¹³⁴ And though O'Brien offers his personal views on a wide range of social and ethical questions he puzzlingly never addresses his own ethical judgment of the morality of committed same-sex relationships.

Other opponents of SSCM are much more upfront about their views on the immorality of all homosexual sexual relationships. O'Brien distances himself from the judgment of John Finnis and Robert George on same-sex relations, but only on the rather narrow grounds that their natural law arguments have irreducibly metaphysical elements that are not apt for public argumentation. Other contemporary natural law theorists—even those who espouse a more metaphysically 'upfront' version of natural law deny that same-sex relationships are necessarily prohibited by natural law (or fail to realize any human good)¹³⁵ and other natural lawyers are highly sympathetic with the case for SSCM.¹³⁶ It might have been helpful for the reader for O'Brien to have situated his understanding of the *moral* issues at stake in considering same-sex relationships in this context.¹³⁷

Where does this leave us in the wider argument about SSCM in the U.S. and elsewhere? Are we stuck in a series of philosophical zero-sum games that can only be resolved by an unambiguous victory by one side or the other in a series of ongoing culture wars? I am not as pessimistic as O'Brien seems to be. Sound argumentation and balanced interpretation of texts can help us at least begin to work through some of the more extreme positions that are present on both sides of the same-sex marriage debate. If we agree with the concept of public reason, in one of its many variations,

¹³⁴ I raise this point as a genuine question rather than as cynical assumption.

¹³⁵ MARK C. MURPHY, *AN ESSAY ON DIVINE AUTHORITY* 176-83 (2002). Murphy is one of the foremost members of a younger generation of natural law philosophers.

¹³⁶ See JEAN PORTER, *MINISTERS OF THE LAW: A NATURAL LAW THEORY OF LEGAL AUTHORITY* 123 (2010); Chartier, *supra* note 94. Some committed lay Roman Catholics in the UK such as Lord Deben (formerly John Gummer, a past Conservative cabinet minister) endorsed the UK Government's same-sex marriage bill (now enacted) on the basis that natural law arguments would support the extension of civil marriage rights to same-sex partners. Lord Deben said "[s]urely our understanding of sexual relationships and sexuality should lead us to understand that there is an extension of natural law from that understanding (i.e. of civil marriage being a natural institution common to humanity). That extension should lead us to be prepared in the state to allow people of the same-sex to marry. It is wrong to suggest that there is something unnatural for them to wish to take this step." H.L. DEB. Dec. 11, 2012, vol 741, col.993.

¹³⁷ Many readers will be surprised by O'Brien's lack of curiosity or appreciation as to how a loving and committed same-sex relationship is categorically different to the relationship between a widower and his brother who may choose to live together for company. Nowhere in the article does O'Brien seek to explore how such same-sex couples may experience distinctive common (human) goods from their relationship in a way that is more similar to the experience of a heterosexual married couple than with two friends or siblings lodging together. Instead O'Brien tendentiously seeks to frame the SSCM question as if it were if it were essentially a dispute about the state's general "endorsement of gay sex". O'Brien, *supra* note 1, at 457.

then we may have some cause to believe that debates in western societies can produce more than just a repeating series of fruitless, agonistic debates. This is not to say that there is only one politically objective response to every vexed ethical-political question, only that in line with our duty of civility to each other, we need to keep working at the answers together.

THE ORIGINATION CLAUSE: MEANING, PRECEDENT, AND THEORY FROM THE 12TH TO 21ST CENTURY

Priscilla H.M. Zotti*

and

Nicholas M. Schmitz**

ABSTRACT

Article 1, Section 7 of the Constitution requires that, “All Bills for raising Revenue shall originate in the House of Representatives; but the Senate may propose or concur with Amendments as on other Bills.” The history of the clause reveals a strong restriction that nobody except the direct representatives of the people familiar with their circumstances can constitutionally propose the laws drawing forth national revenues. Through a handful of cases in the 20th century, the Supreme Court adopted a deferential standard for judging Origination Clause challenges to Senate tax measures. This standard departed from both the original understanding of the Clause and the design of our mixed legislature. The Supreme Court has yet to rule on a large scale Senate tax which significantly challenges its 20th century interpretation. Additionally, it has not articulated a clear and encompassing standard for all potential Senate taxes that might come under challenge on Origination Clause grounds. A historical review of the origins, evolution, and modern interpretation of the constitutional dictate reveals that future challenges to Senate originated taxes may highlight the limits of the 20th century’s permissive standard. Such challenges may force the Court either to modify its standard in favor of finding Origination Clause violations in Senate tax measures or else effectively nullify the Origination Clause requirement ratified in the Constitution. Currently, multiple cases challenging the constitutionality of the Affordable Care Act under the Origination Clause are pending before the Judiciary. If the Court in any one of these challenges does not enforce a germaneness requirement to Senate-originated taxes through amendment, then the Origination Clause will become wholly superfluous. Through focused analysis on the legal tradition of Colonial and State origination requirements, and the balance of evidence from the 1787 Convention and the ratification debates, we find significant historical evidence that such an interpretation of the Clause is not warranted.

CONTENTS

I.	INTRODUCTION	72
II.	HISTORICAL EVOLUTION FROM MAGNA CARTA TO THE COLONIAL PERIOD	75
	A. EARLY COLONIAL EVOLUTION	79

B. THE ENLIGHTENMENT AND REVOLUTIONARY INFLUENCES	81
III. INITIAL STATE CONSTITUTIONS PRIOR TO THE CONSTITUTIONAL CONVENTION OF 1787	85
IV. PHILADELPHIA 1787	91
V. THE MEANING OF WORDS.....	99
A. ORIGINAL PUBLIC UNDERSTANDING AND THE CHANGING JUDICIAL UNDERSTANDING.....	99
i. “Revenue Raising”	99
ii. “Originate”	103
iii. Germaneness and the phrase “as on other Bills”	104
B. FURTHER EVIDENCE FROM THE RATIFICATION DEBATES.....	114
VI. SUPREME COURT PRECEDENT	117
VII. THE COURT’S STANDARD, IMPLICATIONS, AND THEORETICAL CHALLENGES	130
VIII. CONCLUSION	131
APPENDIX	135

I. INTRODUCTION

The patriots of this province ... were, for one hundred and fifty years, allowed to tax themselves, and govern their internal concerns as they thought best. Parliament governed their trade as they thought fit. This plan they wish may continue forever. But it is honestly confessed, rather than become subject to the absolute authority of parliament in all cases of taxation and internal polity, they will be driven to throw off that of regulating trade.

John Adams, 1775¹

*Professor of Political Science, United States Naval Academy Political Science Department, Ph.D., University of Texas.

** MPhil., Oxford University.

The authors thank Professors Erik M. Jensen and Robert G. Natelson for their invaluable criticism and comments on earlier drafts. Above all, the authors recognize Andrew T.

The Origination Clause

In 2012, while striking down as unconstitutional the “State mandate” provision of the Patient Protection and Affordable Care Act, the Supreme Court upheld as constitutional the only other provision of that Act before the Court, the “individual mandate,” which purports to require most Americans to maintain “minimal essential” health insurance coverage. The Court found this provision constitutional under Congress’s power to tax: “In this case, however, it is reasonable to construe what Congress has done as increasing taxes on those who have a certain amount of income, but choose to go without health insurance. Such legislation is within Congress’s power to tax.”² Both opponents and proponents of the law were surprised by the Court’s unexpected reliance on the taxing power of Congress to resolve this landmark case. Most assumed the case would center squarely on the extent of Congress’s power to regulate interstate commerce. For the last century of U.S. jurisprudence, cases and controversies challenging or defending an expansive interpretation of implied federal powers have been argued, won, and occasionally lost on predominantly commerce clause merits. The Court’s unexpected decision calls for renewed scholarly analysis of factors impacting Congress’s taxing power.

This article focuses neither on the merits of the aforementioned case nor on the broader scholarly analysis of Congress’ power to tax. The taxation literature and precedent is already extensive. The focus of this article is confined to one of only a few³ explicit constitutional structures impacting the power of Congress to tax: the Origination Clause of Art.1, Sec.7:

All Bills for raising Revenue shall originate in the House of Representatives; but the Senate may propose or concur with Amendments as on other Bills.

As a specifically procedural restraint, the Origination Clause is unique among the several clauses impacting the constitutional viability of any “money bills.” It is ironic that the Origination Clause has received relatively little scholarly and judicial attention over the past century given the ancient legal origins of the underlying principle, its dominant role in the American Revolution, the degree to which the issue saturated the Convention’s debates in 1787, and the clause’s theoretical implications to the separation of powers within the U.S. federal system. Justice Thurgood Marshall’s 1990 opinion in *U.S. v. Munoz Flores* emphatically reaffirmed the

Hyman’s invaluable insights, comments, revisions, and sharing of sources on the topic. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

¹ John Adams, 8: Novanglus (1775) in *THE REVOLUTIONARY WRITINGS OF JOHN ADAMS, SELECTED AND WITH A FOREWORD BY C. BRADLEY THOMPSON* 245-46 (Indianapolis: Liberty Fund, 2000).

² Nat’l Fed’n. of Ind. Bus. v. Sebelius, 567 U.S. ___, slip op. at 58 (2012).

³ See U.S. CONST. art. I, § 9’s prohibitions on capitation and direct taxes, as well as the prohibition on federal taxes and duties on States’ exports. Additionally, the Sixteenth Amendment’s impact on sources and apportionment is relevant.

Court's "duty to conduct such a review"⁴ of Origination Clause challenges, thereby departing from 90 years of court deference to the legislative branch on the issue when controversies arise. The confluence of these factors as well as the scarcity of scholarly analysis of the judicial status of the Origination Clause post *Munoz-Florez*, makes the issue ripe for renewed scholarly review.

Additionally, the subject has gained significant relevance in the last year as a result of several pending lawsuits challenging the taxes of the Affordable Care Act on Origination Clause grounds. The leading case, *Matthew Sissel v. United States Department of Health and Human Services*, is currently on appeal awaiting oral arguments in District of Columbia Circuit.⁵

Scholarly analysis of the Origination Clause has been exceeding rare.⁶ ^{7 8 9} Much of the research has been narrowly focused on particular court cases, or 20th century Court precedent in general. We know of no thorough, scholarly historical analysis of the origin and intent of the clause, and only two scholarly academic reviews in the aftermath of *Munoz-Florez*

⁴ United States v. Munoz-Florez, 495 U.S. 385, 409 (1990) (Scalia J. concurring).

⁵ Docket Number 13-5202. Another Origination Clause challenge is pending in the United States Court of Appeals for the Fifth Circuit, *Hotze v. Sebelius*, Docket Number 14-20039.

⁶ JOSEPH STORY, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES §870-877 (Boston, Hillard, Gray, and Co. 1833) (hereinafter STORY).

⁷ Noel Sargent, *Bills for Raising Revenue under the Federal and State Constitutions*, 4 MINN. L.REV. 330 (1919-1920). See also Marie T. Farrelly, *Special Assessments and the Origination Clause: A Tax on Crooks?*, 58 FORDHAM L. REV. 447 (1989). Farrelly provides a cursory summary of the history of Court precedent and an extensive analysis the clauses implications for the Victims of Crime Assistance Act of 1984 in the run up to *United States v. Munoz-Florez* (1990). See also Michael Medina, *The Origination Clause in the American Constitution: A Comparative Survey*, 23 TULSA L.J. 165, generally and at 233 (1987). Medina conducts an exhaustive comparative study of both contemporary state and international practices on lower house fiscal prerogatives to conclude that as a result of narrow judicial interpretation/enforcement America "[gives] little actual precedence to revenue bills and accord[s] the more immediate legislative voice of the people less constitutional prerogative than other nations." *Id.* at 233 (hereinafter Sargent, Farrelly, or Medina respectively).

⁸ James V. Saturno, *The Origination Clause of the U.S. Constitution: Interpretation and Enforcement*, (Washington DC, CRS, Mar. 15, 2011). Saturno provides a contemporary and useful primer on the history, Court precedent, and parliamentary precedent surrounding the clause. The publication is particularly useful in exploring the history and mechanics of the House procedure of "blue-slipping" bills it finds objectionable on Origination Clause grounds.

⁹ Ronald J. Krotoszynski, *Reconsidering the Nondelegation Doctrine: Universal Service, the Power to Tax, and the Ratification Doctrine*, 80 INDIANA L.J. 239, (2005). Buried within this piece on a more general topic, Krotoszynski provides the definitive chronological summary of the evolution of the Origination Clause within the Constitutional Convention of 1787 (hereinafter Krotoszynski).

which both narrowly focus on precedent.¹⁰ Likewise, we know of no comprehensive historical analysis. We attempt here to delve a little deeper into the origins and meaning of the clause as it was ratified.

Here we (1) Trace the historical evolution of the legal principle of origination; (2) Detail its development in the Constitutional Convention; (3) Analyze what the words meant to those who ratified it, and (4) Review significant Court precedent through *Munoz-Florez* that are relevant to current and future Origination Clause challenges.

From our analysis, this article concludes that throughout the 20th century the Court has developed a historically narrow standard for what bills are considered “Bills for raising Revenue” within the context of Art. 1, §7, and, if classed as a revenue raising bill, a standard that any Senate amendments must be germane to the subject matter of the House originated bill. While the somewhat passive evolution of this standard over the 20th century has survived as relatively uncontroversial given the small scale and the nature of the cases presented, the Court will have to revisit the standard if broader challenges are presented in order to preserve any substantive meaning and effect in the Origination Clause of the Constitution and our theory of mixed legislatures. In the absence of judicial review, the “Aristocratic Branch” of the federal legislature may continue to institutionalize creative legislative maneuvers for originating broader and more burdensome taxing measures in contravention of the Framers’ fear that “If the Senate can originate, they will in the recess of the Legislative Sessions, hatch their mischievous projects, for their own purposes, and have their money bills ready cut & dried, (to use a common phrase) for the meeting of the H. of Representatives.”¹¹

II. HISTORICAL EVOLUTION: FROM MAGNA CARTA TO THE COLONIAL PERIOD

And just as the charter was claimed by the English Radicals as a natural birthright, so in America some of the principles came to be established as individual rights enforceable against authority in all its forms, whether legislative, executive or judicial ... Crown, governor or council, or later by state and federal government.”¹²

The legal influences on the American Origination Clause dates back to at least the 1215 AD *Magna Carta* forced upon King John at Runnymede by his Barons following their open insurrection against the Crown. Earlier influence from the British Constitutional tradition may be attribut-

¹⁰ Rebecca Kysar, *The ‘Shell Bill’ Game: Avoidance and the Origination Clause*, Brooklyn Law School Legal Studies Research Papers, 32 (May 2013) (hereinafter Kysar); see also Timothy Sandefur, *So It’s A Tax, Now What? Some of the Problems Remaining After NFIB V. Sebelius*, 17 TEXAS REV. L. & POLITICS 203 (2013).

¹¹ JAMES MADISON, NOTES ON DEBATES IN THE FEDERAL CONVENTION OF 1787, 443 (New York, Norton & Co. Inc., 1969) (hereinafter MADISON).

¹² J.C. HOLT, MAGNA CARTA 17 (2d ed. 1992) (hereinafter HOLT).

ed to the the much more succinct 1100 AD “Charter of Liberties of Henry I” in that the principle of the “common counsel” of the king was invoked as justification for the Crown’s power in two of its clauses. However, the link between forms of taxation specifically and some degree of “popular” control was not made nearly as explicit in the 1100 Charter as in the 1215 *Magna Carta*. Of the 63 clauses in the 1215 *Magna Carta* two in particular form the constitutional genesis of the codification of origination limitations: chapters 12 and 14.¹³ Chapter 12 specifies that “*No scutage or aid is to be levied in our realm except by the common counsel of our realm. . . .*” Chapter 14 extensively details who composed the “common counsel of the realm,” procedural restrictions on when and how they were to be convened, and what constituted their consent.¹⁴ At their cores, chapter 12 established the principle while chapter 14 specified the procedures and conditions of transparency thought necessary to safeguard the principle.¹⁵

It would be misleading, however, to reduce chapters 12 and 14 of the *Magna Carta* to a continuous strand of legal precedent inherited by the American colonists and reaffirmed in the U.S. Constitution. What was codified in 1215 was far from the absolute popular prerogative against general taxation that animated the maxim in the American Revolution against “taxation without representation.” The 1215 clauses are less ambitious in several key respects, as discussed below.

First, “scutage” and “aid” were two narrow forms of taxation levied by the Crown in the 12th and 13th centuries. Scutage was a fee paid to the Crown in exchange for a release from military obligations in various military campaigns for which scutage was levied. Aid was a general term for various forms of feudal fees provided to the lord or Crown on occasions such as marriages, knightings, and ransoms. Far from limiting a general, centralized power of taxation without popular consent, chapter 12 aimed to deny the crown of these two specific forms of monetary extraction that had been particularly abused under the reign of King John:

¹³ The original Latin document was continuous without indexing. We use the common convention.

¹⁴ “And to obtain the common counsel of the realm for the assessment of an aid (except in the cases aforesaid) or scutage, we will have archbishops, bishops, abbots, earls, and greater barons summoned individually by our letters, and we shall have summoned generally through our sheriffs and bailiffs all those who hold of us in chief, for a fixed date, with at least forty days’ notice, and at a fixed place; and in all letters of summons we will state the reason for the summons. And when the summons has thus been made, the business shall go forward on the day arranged according to the counsel of those present, even if not all those summoned have come.” See HOLT, *supra* note 12, appendix 6, at 455.

¹⁵ WILLIAM SHARP MCKECHNIE, *MAGNA CARTA: A COMMENTARY ON THE GREAT CHARTER OF KING JOHN*, Ch. 14 (1914) (hereinafter MCKECHNIE): “This chapter, which has no equivalent among the Articles of the Barons, appears here incidentally: it would never have found a place in *Magna Carta* but for the need of machinery to give effect to chapter 12.”

The Origination Clause

It is a commonplace of our text-books that chapters 12 and 14, taken together, amount to the Crown's absolute surrender of all powers of arbitrary taxation, and even that they enunciate a doctrine of the nation's right to tax itself. Yet the very idea of "taxation" in its abstract form, as opposed to specific tallages and exactions, levied on definite things or individuals, is essentially modern. . . . A regular scheme of "taxation" to meet the ordinary expenses of government was undreamt of. It is too much to suppose, then, that our ancestors in 1215 sought to abolish something which, strictly speaking, did not exist. The famous clause treats, not of "taxation" in the abstract, but of the scutages and aids already discussed.¹⁶

Second, the principle of popular consent present in chapters 12 and 14 of *Magna Carta* was far less egalitarian than its ideological reincarnations in the 17th and 18th centuries. The "common counsel" in 13th century England required to consent to these forms of taxation was aristocratic to its core by modern standards.

Finally, it is important to note that chapters 12 and 14 were removed from *Magna Carta* in its subsequent reissuing by the Crown in 1216, 1217, and 1225. One explanation offered by historians for this omission is that most of *Magna Carta* was actually a reaffirmation of ancient customs and privileges afforded to the Barons and clergy by the Crown. Far from being a wholly revolutionary document, most of the *Magna Carta* of 1215 was a forced confirmation by the Crown of abused privilege and custom, with the exception that chapter 12's requirement on scutage "had no legal basis."¹⁷

Even though it was removed in the reissued Charters, the idea and custom of obtaining popular consent through strict procedures before royal taxation remained rooted in the courts and counsel of the Crown. This was evident in Parliament's early refusal of various royal exactions in 1242 and 1255 on the grounds that the counsel's consent had not met the prerequisites in the 1215 *Magna Carta*.¹⁸

¹⁶ *Id.* Ch. 12.

¹⁷ HOLT, *supra* note 12, at 318. See also *id.* at 301 ("here [scutage] the Charter stated not law but innovation" and at 317 "In the matter of aids, it simply reasserted the usual process of consent.") See also MCKECHNIE, *supra* note 15, chapter 12: "The total omission of this chapter in 1216 may have been partly occasioned by the consciousness that it contained an innovation unwarranted by custom: the reissue of 1217 said nothing of aids, and contented itself, in regard to the vexed question of scutages, with the vague declaration that for the future these should be taken as had been the custom under Henry II. In spite, however, of the omission of chapter 12 from all reissues of the Great Charter, it was customary for Henry's advisers to consult "the Common Council" before exacting a scutage or aid. This was done, for example, in 1222, when a Council granted an "aid for the Holy Land" of three marks for an earl, one mark for a baron, and twelve pence for a knight. The consent of a Council, indeed, was usually taken even for one of the three recognized feudal aids."

¹⁷ MCKECHNIE, *supra* note 15, ch. 14.

¹⁸ *Id.*

None of these limitations of *Magna Carta* are meant to diminish its historical significance or the credit due to the document in influencing subsequent constitutional structures codified in England and across the Atlantic. However, it is important to note that the historical context of chapters 12 and 14 are quite different from the ideals and principles debated in the formation of the British Bill of Rights and the Constitutional Convention in Philadelphia. How then did the custom proceed from 1215 to William Penn's animated conception that "[n]o Law can be made, no Money Levied, nor a Penny Legally Demanded (even to defray the Charges of the Government) without your own Consent"?¹⁹

By Richard II's reign (1377-1399) it was customary that the "Commons granted with the assent of the Lords."²⁰ The principle remained for several hundred years, but was not firmly solidified against the claims of the Lords until the late 17th century. In 1671 a battle between the Commons and the Lords erupted when the Lords attempted to reduce a tax on sugar that the Commons had originated. The Lords recognized the principle that the Commons exclusively originate new taxes, but the Lords reasoned in this case that they were reducing revenue vice raising it. On July 3rd, 1678, the Commons passed a resolution that the Lords had no power to amend revenue measures. The Lords fought the Commons on this minor prerogative of at least reducing revenue until the 1690s when the Commons effectively won the exclusive right to manage all revenues.²¹ It is notable that the prerogative the colonists brought with them from the British Parliament of the late 17th century included a lower House right to "all bills for purpose of taxation, or containing clauses imposing a tax."²² This is a far broader category of legislation than the 20th century Ameri-

¹⁹ WILLIAM PENN, "England's Great Interest, in the Choice of this New Parliament Dedicated to All Her Free-Holders and Electors" (1679). Additionally see Penn's later instructional text to the Colonists, "The Excellent Priviledge of Liberty and Property Being the Birth-Right of the Free-born Subjects of England," 1687: "In England the Law is both the measure and the bound of every Subject's duty and allegiance, each man having a fixed Fundamental Right born with him, as to freedom of his person and property in his estate, which he cannot be deprived of, but either by his consent, or some crime, for which the law has imposed such a penalty or forfeiture."

²⁰ ROBERT LUCE, LEGISLATIVE PROBLEMS: DEVELOPMENT, STATUS, AND TREND OF THE TREATMENT AND EXERCISE OF LAWMAKING POWERS, 390 (1935) (hereinafter LUCE). For earlier assertions by Parliament in general of this taxation prerogative, see the 1627 "Petition of Right" against Charles I, 3 Chas.1 c.1 §8:

"[T]hat no man hereafter be compelled to make or yield any gift loan benevolence tax or such like charge without common consent by Act of Parliament, and that none be called to make answer or take such oath or to give attendance or be confined or otherwise molested or disquieted concerning the same or for refusal thereof."

²¹ LUCE, *supra* note 20, at 390. See also Sargent, *supra* note 7, at 334: "In the British Parliament, in 1678, it was settled that: (1) 'all bills for purpose of taxation, or containing clauses imposing a tax, must originate in the House of Commons and not in the House of Lords.'"

²² *Id.*

can courts' concept of "incidental taxation" outside of the scope of "revenue raising bills."

A. EARLY COLONIAL EVOLUTION

The colonial history of popular procedural limitations on taxation is mixed. The royal charters issued during the 17th century have various degrees of restraint specified. Generally, charters granted before the 1660s have little popular involvement required by the charters' language. Charters granted in the latter half of the 17th century have more robust requirements and language with respect to taxation.

Colonial charter's granted during the first half of the 17th century under King James I and Charles I, generally afforded colonial governors broader and less popularly constrained methods of taxation. The Maryland charter of 1632 granted the Barron of Baltimore and his heirs power "to assess and impose the said Taxes and Subsidies" on the colonists subject only to the limitation that they be "reasonably assessed", and "upon just Cause and in due Proportion."²³ Likewise, the 1629 colonial charter of Massachusetts issued to the "Councell established at Plymouth" placed no popular restraint on taxation.²⁴

Charters granted after the restoration of the House of Stuart between the 1660s and 1690s generally mandated some form of local, popular consent for taxation. For example, the **Carolina** charter of 1663 gave power to make and enact taxes provided the "advice, assent and approbation of the freemen of the said province, or the greater part of them, or of their delegates or deputies."²⁵

The 1681 charter for Pennsylvania granted to William Penn used almost identical language and further required of "Laws . . . for the raising of money for the publick use of the said Province" the "advice, assent, and approbation of the Freemen of the said Countrey, or the greater parse of them, or of their Delegates or Deputies."²⁶ In the subsequent 1683 "Frame of Government in Pennsylvania," the constitution established that "no

²³ The Charter of Maryland: 1632, available at

http://avalon.law.yale.edu/17th_century/ma01.asp.

²⁴ The Charter of Massachusetts Bay : 1629 available at

http://avalon.law.yale.edu/17th_century/mass03.asp.

²⁵ The Charter of Carolina - 1663, *reprinted in* THE FEDERAL AND STATE CONSTITUTIONS, COLONIAL CHARTERS, AND OTHER ORGANIC LAWS OF THE UNITED STATES PART II, at 1384 (Washington, GPO, 2d ed. 1878) (hereinafter GPO). Interestingly, authority for legislative ordinance by the less popular direction of the assembly of eight specified in the charter was authorized "because such assemblies of freeholders cannot [always] be conveniently called." However, such legislative authority was limited from "extend[ing] to the binding, charging, or taking away of the right or interest of any person or persons, in their freehold, goods or chattels whatsoever." *Id.* at 1384-85. The 1663 Carolina charter also required on commerce taxation specifically that "the said customs [are] to be reasonably assessed, upon any occasion, by themselves, and by and with the consent of the free people there, or the greater part of them as aforesaid." *Id.* at 1510.

²⁶ Charter for the Province of Pennsylvania, *reprinted in* GPO, *supra* note 25, at 1510.

money or goods, shall be raised upon, or paid by, any of the people of the province by way of public tax, custom or contribution, but by law . . . ; and whoever shall levy, collect, or pay any money or goods contrary thereunto, shall be held a public enemy to the province and a betrayer of the liberties of the people thereof.”²⁷ Furthermore, the modern difficulties of resolving the question of what constitutes a revenue raising bill would have been unproblematic under Penn’s Frame of Government as it required that “not taxes should be levied but by a law for that purpose made.”²⁸

In 1688, James II gave New England “full power and authority by and with the advise and consent of our said Councill, or the major part of them, to impose assess and raise and levy rates and taxes as you shall find necessary”²⁹

The 1691 Massachusetts charter added a requirement for the “advice and Consent of the Councill” in its grant of power “to impose and leavy proportionable and reasonable Assessment Rates and Taxes. . . .”³⁰

However, even when origination or popular consent requirements were not mandated in the royal charters, many colonies simply wrote popular assembly origination requirements into their own governing laws. For example, Maryland passed a law binding its upper council and governor in 1650 entitled “An ACT against raising of Money within this Province, without Consent of the Assembly.” The law required:

That no Subsidies, Aids, Customs, Taxes or Impositions, shall hereafter be laid, assessed, levied or imposed, upon the Freemen of this Province, or on their Merchandize, Goods or Chattels, without the Consent and Approbation of the Freemen of this Province, their Deputies, or the major Part of them, first had and declared in a General Assembly of this Province.³¹

This most early example of an origination requirement in the American colonies substitutes the more variable verb “originate” used in our Constitution with the more explicit specification that taxes must be “first had and declared in General Assembly.” The law was likely introduced by Lord Baltimore to counter his critics and assuage the freemen of the colony

²⁷ Frame of Government of Pennsylvania – 1683, *reprinted in* GPO, *supra* note 25, at 1524.

²⁸ 1 STORY, *supra* note 6, §§ 112, 123.

²⁹ Commission of Sir Edmund Andros for the Dominion of New England, Apr. 7, 1688 available at http://avalon.law.yale.edu/17th_century/mass06.asp.

³⁰ The Charter of Massachusetts Bay – 1691 available at http://avalon.law.yale.edu/17th_century/mass07.asp.

³¹ FRANCIS BACON, THE LAWS OF MARYLAND, VOL. 75, 37-38 (1765) available at *Archives of Maryland Online*, <http://aomol.net/megafile/msa/speccol/sc2900/sc2908/000001/000075/html/index.html>.

that the method of their taxation would be void of any arbitrary, unpopular influence.³²

New Jersey likewise wrote its own strict requirement into its laws in 1681: “That it shall not be lawful . . . to levy or raise any sum or sums of money, or any other tax whatsoever, without the act, consent and concurrence of the General Free Assembly.”³³

B. The Enlightenment and Revolutionary Influences

The liberalization of colonial charters generally paralleled shifts in political thought across the Atlantic. With the execution of King Charles I in 1649 and the deposing of James II in the late 1680s, the supremacy of popular rule was firmly established in the minds of British subjects and her colonists in the Americas. The British Bill of Rights of 1689 following the Glorious Revolution mandated, “[t]hat levying money for or to the use of the crown, by pretence of prerogative, without grant of parliament, for longer time, or in other manner than the same is or shall be granted, is illegal.”³⁴

On the American continent the colonists experimented with representative government throughout the 17th and 18th centuries. Most colonial charters established only a royal governor and council subject to the Crown’s influence. However, most of the colonies by the 18th century had instituted popularly elected lower houses similar to the House of Commons in Parliament. Under the various names of “House of Delegates,” “Burgesses,” “Commons,” or “Representatives,” these lower, popularly elected chambers were often given unique functions, privileges, and powers distinct from the colonial councils.

By 1776, many States had a lower-house with some advantage over the upper-house on monetary and taxing matters either by constitutional mandate, statute, or common practice. While origination restrictions were common in colonial legislatures prior to the revolutionary period, the primary revolutionary grievance of unprecedented, distant taxation measures without their local consent through Parliament’s Sugar and Stamp Acts

³² JOHN L. BOZMAN, *THE HISTORY OF MARYLAND FROM ITS FIRST SETTLEMENT IN 1633 TO THE RESTORATION IN 1660*, VOL. 2, 401 (1837). *Also see* TIMOTHY RIORDAN, *THE PLUNDERING TIME: MARYLAND AND THE ENGLISH CIVIL WAR*, 316-25; 327-30 (Baltimore, Maryland Historical Society, 2004): Following a period of open insurrection by Protestants in Maryland against Lord Baltimore’s Catholic government due in part to the spillover from the English Civil War, Lord Baltimore was under significant pressure to curry favor with the victorious Parliamentarians to sustain his rule as a Catholic proprietor. Among the many liberalizing policies he instituted between 1649 and 1650 were the establishing a bicameral legislature in Annapolis, the passing of the continent’s second religious toleration act, and restrictions on taxation to the newly created popular assembly.

³³ *THE GRANTS, CONCESSIONS, AND ORIGINAL CONSTITUTIONS OF THE PROVINCE OF NEW JERSEY* 424 (Aaron Learning & Jacob Spicer eds., 1881).

³⁴ 1 Will. & Mary, Sess.2 c.2. § 4; British Bill of Rights § 4. (1688).

solidified the ideological convictions and constitutional structures in the various States after independence.

The advent of British social contract theory between Hobbes and Locke in 1651 and 1689 respectively with their emphasis on consent theory and property undoubtedly molded the Enlightenment reincarnation of origination requirements on both continents. For the colonists the objections in the late 18th century appealed to a curious mix of as much custom, common law, and privilege as to natural law and the deontological ethics of universal, inherent rights.

Prior to the 1760s the colonists had enjoyed not only the privilege of local ratification of any proposed taxing measures by the Crown and Parliament but more often than not the original design of the taxing measures themselves. As the Barons at Runnymede had become accustomed to the royal privilege and custom of ratifying any new aids levied to the Crown, so too did the colonists feel it had become their prerogative to design and approve of any new internal taxation on the Colonies. With the colonists, the case was even more explicit as many of the previously cited royal charters granted that authority while reserving more power to the Crown on matters of regulating colonial exports and commerce.³⁵

With reverence for the prerogative, William Pitt (“the elder”) protested in Parliament in 1765 on behalf of the colonists against the Stamp Act by arguing that the “distinction between legislation and taxation is essentially necessary to liberty. . . . The Commons of America, represented in their several assemblies, have ever been in possession of the exercise of this their constitutional right of giving and granting their own money. They would have been slaves if they had not enjoyed it.”³⁶

Many colonists likewise disagreed with the constitutional logic of Parliamentary supremacy in taxation when its effects materialized. In protest of the Sugar Act of 1764 the Virginia House of Burgesses (along with separate petitions from ten other colonies) sent its famous petition to the House of Commons citing the colonial logic of opposition to the internal tax:

[T]he Council and Burgesses . . . in a respectful manner but with decent firmness, to remonstrate against such a measure . . . conceive it is essential to British liberty that laws imposing taxes on the people ought not to be

³⁵ See WILLIAM DOUGLASS, A SUMMARY, HISTORICAL AND POLITICAL, OF THE FIRST PLANTING, PROGRESSIVE IMPROVEMENTS, AND PRESENT STATE OF THE BRITISH SETTLEMENTS IN NORTH AMERICA, 212 (1748) (hereinafter DOUGLASS) (“The vacating of all charter and proprietary governments is not the ultimate chastisement that may be used with delinquent colonies; the parliament of Great Britain may abridge them of many valuable privileges which they enjoy at present; . . . therefore the colonies ought to be circumspect, and not offend their mother-country; as for instance 1. In abusing that privilege which our colonies have in ratifying taxes and affecting of themselves;”).

³⁶ William Pitt, *On an address to the Throne, in which the right of taxing America is discussed*, Dec. 17, 1765 reprinted in THE TREASURY OF BRITISH ELOQUENCE 140-41 (Robert Cochrane, ed., London & Edinburgh, 1877).

The Origination Clause

made without the consent of representatives chosen by themselves; who, at the same time that they are acquainted with the circumstances of their constituents, sustain a proportion of the burden laid on them.³⁷

The logic was echoed in the fundamental objection of the first act of coordinated American government in the Stamp Act Congress:

3d. That it is inseparably essential to the freedom of a people, and the undoubted rights of Englishmen, that no taxes should be imposed on them, but with their own consent, given personally, or by their representatives. .

. .

5d. That the only representatives of the people of these colonies, are persons chosen therein by themselves; and that no taxes ever have been, or can be constitutionally imposed on them, but by their respective legislatures.

The First Continental Congress in October of 1774 reiterated the same philosophy in their declaration of colonial rights and grievances. The opening sentence of the declaration states,

Whereas, since the close of the last war, the British Parliament, claiming a power, of right, to bind the people of America by statutes in all cases whatsoever, hath in some acts expressly *imposed* taxes on them, and in others, under various pretenses, but in fact for the purpose of raising revenue, hath imposed rates and duties payable in these Colonies, established a Board of Commissioners, *with unconstitutional powers*... .

Of particular relevance here to contemporary jurisprudence on Origination Clause issues is the colonists' insistence (theoretically justified or not) on semantically reducing Parliamentary taxation measures passed under "various [legislative] pretenses" to "revenue raising bills." There is a logical case that the various Parliamentary exactions of the 1760s and 1770s could easily have been construed as acts directly to fund/reimburse narrow and constitutional legislative purposes. Such an interpretation would make revenue raising "merely incidental" to the legislative purpose of providing for the local defense of the colonies. While the Sugar Act was known as a revenue raising act, Parliament's position was that it was "for defraying the expenses of defending, protecting, and securing the [colonies]." Likewise, the Stamp Act was justified under the specific purpose of reimbursing the British government for the local defense expenses it had burdened in support of the colonies during the Seven Years War. Additionally, it made little difference to the colonists that Prime Minister Grenville had solicited proposals from various colonial representatives (to include Benjamin Franklin) and MP's on the tax prior to its institutions. The argument for "virtual representation" was flawed in the colonists' minds as they insisted repeatedly in their grievances on their local representation

³⁷ Virginia House of Burgesses, *Petition of the Virginia House of Burgesses to the House of Commons*, (Dec. 18, 1764) available at http://avalon.law.yale.edu/18th_century/petition_va_1764.asp.

familiar with the circumstances of their constituents. The colonists clearly construed (whether logically or illogically) legislative and statutory ambiguities over what constituted a revenue raising bills in favor of finding violations of Origination Clause principles and constitutional guarantees.

The declaration of the First Continental Congress goes on to elaborate on the constitutional philosophy backing their grievance:

That the foundation of English liberty, and of all free government, is a right in the people to participate in their legislative council; and as the English colonists are not represented, and from their local and other circumstances cannot be properly represented in the British Parliament, they are entitled to a free and exclusive power of legislation in their several Provincial Legislatures, where their right of representation can alone be preserved, in all cases of taxation and internal polity, subject only to the negative of their sovereign, in such a manner as has been heretofore used and accustomed.

Interestingly, the original draft of the document (before stylization and adoption) included the additional statement,

That all the statutes before mentioned, for the purpose of raising a revenue, by imposing ‘rates and duties’ payable in these Colonies, establishing a Board of Commissioners, and extending the jurisdiction of Courts of Admiralty, for the collection of such ‘rates and duties,’ are illegal and void.³⁸

John Adams in 1775 in Boston likewise dissented but in less tactful language. He vehemently disagreed with the constitutional authority of unrepresented taxation, and warned that the consequence for Parliament infringing on the colonist’s right of self-taxation was not merely that the colonists would insist on its restoration, but also that Britain would lose even its legitimate sovereign authority to regulate colonial commerce as well:

That there are any who pant after “independence,” (meaning by this word a new plan of government over all America, unconnected with the crown of England, or meaning by it an exemption from the power of parliament to regulate trade,) is as great a slander upon the province as ever was committed to writing. The patriots of this province desire nothing new; they wish only to keep their old privileges. They were, for one hundred and fifty years, allowed to tax themselves, and govern their internal concerns as they thought best. Parliament governed their trade as they thought fit. This plan they wish may continue forever. But it is honestly confessed, rather than become subject to the absolute authority of parlia-

³⁸ First Continental Congress, *Original Draught of the Declaration of Rights and of Grievances, Made by the Congress of 1774*, in *THE WORKS OF JOHN ADAMS VOL. II*, 538 Appendix C (1850).

ment in all cases of taxation and internal polity, they will be driven to throw off that of regulating trade.³⁹

The rest is history. However, the primacy of this Revolutionary era ideological cause against unprecedented taxation measures procedurally implemented by distant counsels unfamiliar with the “circumstances of their constituents” impacted the constitutional structures of the early State constitutions with respect to revenue origination requirements.

III. INITIAL STATE CONSTITUTIONS PRIOR TO THE CONSTITUTIONAL CONVENTION OF 1787

A cursory textual survey of the various States’ origination requirements between the Revolution and the ratification of the national Constitution are listed as follows:⁴⁰

- Delaware’s constitution required that,

All money-bills for the support of government shall originate in the house of assembly [lower house], and may be altered, amended, or rejected by the legislative council. All other bills and ordinances may take rise in the house of assembly or legislative council, and may be altered, amended, or rejected by either.⁴¹

Delaware’s rewritten constitution of 1792 accomplished an interesting modification of the Origination Clause similar to Maryland’s 1776 constitution:

All bills for raising revenue shall originate in the house of representatives; but the senate may propose alterations, as on other bills; and no bill, for the operation of which, when passed into a law, revenue may incidentally arise, shall be accounted a bill for raising revenue; nor shall any matter or clause whatever, not immediately relating to and necessary for raising revenue, be in any manner blended with or annexed to a bill for raising revenue.⁴²

- Georgia’s 1777 constitution established a unicameral legislature and thus had no need for origination restrictions. The State’s 1789 constitution created a bicameral legislature and the 1798 constitu-

³⁹ JOHN ADAMS, 8: *Novanglus*, (1775) in *THE REVOLUTIONARY WRITINGS OF JOHN ADAMS*, SELECTED AND WITH A FOREWORD BY C. BRADLEY THOMPSON 245-46 (Indianapolis: Liberty Fund, 2000).

⁴⁰ Of the 13 Original States, Connecticut and Rhode Island continued their charter’s and common law practices as the organic laws of the state between the revolution and the ratification of the national constitution. Connecticut drafted a constitution in 1818 that did not include origination restrictions, and Rhode Island drafted its constitution in 1842 without origination restrictions.

⁴¹ Delaware Constitution of 1776, *reprinted in* GPO, *supra* note 25, at 1274.

⁴² Delaware Constitution of 1792, *reprinted in* GPO, *supra* note 25, at 1281.

tion added the origination restriction that “All bills for raising revenue or appropriating moneys shall originate in the house of representatives, but the senate shall propose or concur with amendments, as in other bills.”⁴³

- Maryland’s constitution had one of the most instructive and nuanced origination requirements:

X. That the House of Delegates may originate all money bills, propose bills to the Senate, or receive those offered by that body; and assent, dissent, or propose amendments. . . .

XI. That the Senate may be at full and perfect liberty to exercise their judgment in passing laws-and that they may not be compelled by the House of Delegates, either to reject a money bill, which the emergency of affairs may require, or to assent to some other act of legislation, in their conscience and judgment injurious to the public welfare--the House of Delegates shall not on any occasion, or under any presence annex to, or blend with a money bill, any matter, clause, or thing, not immediately relating to, and necessary for the imposing, assessing, levying, or applying the taxes or supplies, to be raised for the of government, or the current expenses of the State: and to prevent altercation about such bills, it is declared, that no bill, imposing duties or customs for the mere regulation of commerce, or inflicting fines for the reformation of morals, or to enforce the execution of the laws, by which an incidental revenue may arise, shall be accounted a money bill: but every bill, assessing, levying, or applying taxes or supplies, for the support of government, or the current expenses of the State, or appropriating money in the treasury, shall be deemed a money bill.

XII. That no aid, charge, tax, fee, or fees, ought to be set, rated, or levied, under any presence, without consent of the Legislature.⁴⁴

- The Massachusetts constitution of 1780 stipulated that,

[N]o subsidy, charge, tax, impost, or duties, ought to be established, fixed, laid, or levied, under any pretext whatsoever, without the consent of the people, or their representatives in the legislature. . . . [and] all money-bills shall originate in the House of Representatives; but the Senate may propose or concur with amendments, as on other bills.⁴⁵

Massachusetts likewise included similar language to other State’s constitutions in the same document’s declaration of rights: “no part of the property of any individual can, with justice, be taken

⁴³ Georgia Constitution of 1798, *reprinted in* GPO, *supra* note 25, at 389.

⁴⁴ Maryland Constitution of 1776, *reprinted in* GPO, *supra* note 25, at 822.

⁴⁵ Massachusetts Constitution of 1780, *reprinted in* GPO, *supra* note 25, at 959, 964.

from him, or applied to public uses, without his own consent, or that of the representative body of the people.”⁴⁶

This was no significant change for Massachusetts as the practice had been mandated in its colonial legislature: “The house of Representatives is fit upon several privileges . . . 2. That the council [upper house] may only concur or not concur. A tax or any other money-bill, but may make not amendment; the affair of supplying the treasury always originates in the House of Representatives.”⁴⁷

- New Hampshire’s one page Constitution of 1776 required “That all bills, resolves, or votes for raising, levying and collecting money originate in the house of Representatives.”⁴⁸ The more elaborate 1784 constitution specified that “. . . no part of man’s property shall be taken from him, or applied to public uses, without his own consent, or that of the representative body of the people.”⁴⁹ The same document, several clauses later, somewhat redundantly claimed that “No subsidy, charge, tax, impost or duty shall be established, fixed, laid or levied, under any pretext whatsoever, without the consent of the people or their representatives in the legislature, or authority derived from that body.”⁵⁰ In outlining the specific separation of legislative powers, the 1784 constitution mandated that “All money bills shall originate in the house of representatives, but the senate may propose or concur with amendments as on other bills.”⁵¹
- New York’s constitution of 1777 established a bicameral legislature with indirect (by the lower house) elections of the upper house members. However, the Senate was generally proportionally representative of the populace based on a reoccurring census. The State had no explicit origination requirement in its 1777 constitution. The State’s 1821 constitution explicitly clarified that there was no origination restriction in the State legislature: “Any bill may originate in either house of the legislature; and all bills passed by one house may be amended by the other.”⁵²
- North Carolina’s constitution required that “the people of this state ought not to be taxed, or made subject to the repayment of any impost or duty, without the consent of themselves, or their

⁴⁶ *Id.* at 958.

⁴⁷ DOUGLASS, *supra* note 35, at 492-93.

⁴⁸ Constitution of New Hampshire – 1776, *reprinted in* GPO, *supra* note 25, at 1280.

⁴⁹ New Hampshire Constitution of 1784, *reprinted in* GPO, *supra* note 25, at 1281.

⁵⁰ *Id.* at 1283.

⁵¹ *Id.* at 1287.

⁵² New York Constitution of 1821, *reprinted in* GPO, *supra* note 25, at 1342.

Representatives in General Assembly, freely given.”⁵³ The constitution established a bicameral legislature, but did not privilege the lower house in taxation origination. However, both houses were annually elected by the people. Representation was weighted generally equally among the State’s counties in both houses.

- Pennsylvania’s constitution of 1776 vested all legislative power in one popularly elected “House of Representatives of the freemen of the commonwealth.” Therefore, an origination restriction against an upper chamber would have been pointless. However, the constitution specified that “no part of a man’s property can be justly taken from him, or applied to public uses, without his consent, or that of his legal representatives.”⁵⁴ The constitution also required that:

No public tax, custom or contribution shall be imposed, or paid by the people of this state, except by a law for that purpose; And before any law be made for raising it, the purpose for which any tax is to be raised ought to appear clearly to the legislature to be of more service to the community than the money would be, if not collected; which being well observed, taxes can never be burthens.⁵⁵

In 1790 when Pennsylvania rewrote its constitution with legislative power divided between an upper and lower house, it added the origination restriction that “All bills for raising revenue shall originate in the house of representatives; but the senate may propose amendments, as in other bills.”⁵⁶

- South Carolina’s origination requirement in its 1776 Constitutions required that:

All money-bills for the support of government shall originate in the general assembly [lower-house], and shall not be altered or amended by the legislative council, but may be rejected by them. All other bills and ordinances may take rise in the general assembly or legislative council, and may be altered, amended, or rejected by either.⁵⁷

The “legislative council” was elected not popularly, but by vote within the general assembly. The same section of the State’s 1778 version retains the exact same mechanism while changing the chambers’ names to “House of Representatives” and “Senate” and adding the requirement “that no money be drawn out of the public treasury but by the legislative authority of the State.”⁵⁸ In the

⁵³ Constitution of North Carolina – 1776 *reprinted in* GPO, *supra* note 25, at 1410.

⁵⁴ Pennsylvania Constitution of 1776, *reprinted in* GPO, *supra* note 25, at 1541.

⁵⁵ *Id.* at 1547.

⁵⁶ Pennsylvania Constitution of 1790, *reprinted in* GPO, *supra* note 25, at 1550.

⁵⁷ Constitution of South Carolina – 1776, *reprinted in* GPO, *supra* note 25, at 1617.

⁵⁸ Constitution of South Carolina – 1778, *reprinted in* GPO, *supra* note 25, at 1623-1624.

The Origination Clause

1790 Constitution, they kept the lower-house origination requirement, but allowed that revenue raising bills “may be altered, amended, or rejected by the senate.”⁵⁹

- Vermont’s constitution of 1777 stipulated that “no part of a man’s property can be justly taken from him, or applied to public uses, without his consent, or that of his legal representatives,” and that “[a]ll fines, licence money, fees and forfeitures, shall be paid, according to the direction hereafter to be made by the General Assembly.”⁶⁰ The legislative power was unicameral in the 1777 constitution and therefore, an origination restriction would be pointless. After legislative authority was vested in an upper house, the 1863 amendment to the Vermont Constitution added the following origination requirement: “That all revenue raising bills shall originate in the house of representatives, but the senate may propose or concur with amendments, as on other bills.”⁶¹
- Virginia’s constitution had exclusive origination authority for ALL legislation in its lower house, and further barred any amendments to money-bills in its upper house:

All laws shall originate in the House of Delegates, to be approved of or rejected by the Senate, or to be amended, with consent of the House of Delegates; except money-bills, which in no instance shall be altered by the Senate, but wholly approved or rejected.⁶²

Of the eleven available State constitutions immediately following the American Revolution, eight established bicameral legislatures (nine by 1790 with PA, and 10 by 1863 with VT). Of those nine with bicameral legislatures by 1790, seven had lower house Origination Clauses (NY and NC had no Origination Clause; however, North Carolina had annual senatorial elections). Of the seven with Origination Clauses by 1790, six allowed upper-house amendments to revenue raising bills (VA prohibited senate amendments.⁶³ SC had amended this in 1790 to allow senate amendments). On Origination Clause codification practices leading up to the ratification of the national Constitution, New York and Virginia repre-

⁵⁹ Constitution of South Carolina – 1790, *reprinted in* GPO, *supra* note 25, at 1630.

⁶⁰ Vermont Constitution of 1777, *reprinted in* GPO, *supra* note 25, at 1860, 1864.

⁶¹ Vermont Amendments to the Constitution of 1793, Art. III *reprinted in* GPO, *supra* note 25, at 1836, 1883.

⁶² Virginia Constitution of 1776, *reprinted in* GPO, *supra* note 25, at 1909-10.

⁶³ Although, on Maryland’s it may be argued that “From a provision aimed at riders it may be inferred that the Maryland Senate could not originally amend a money bill; in 1851 the power of either branch to amend any measure was definitely specified.” *See* LUCE, *supra* note 20, at 415. For arguments sake we class Maryland’s constitution as not explicitly forbidding senate amendments.

sented outliers at opposite extremes given the common practice of requiring that revenue raising bills originate in the lower house while allowing the upper house to amend such bills. Additionally, it is relevant to the Court's 20th century interpretation of Senate bills that incidentally raise revenue, that the only two State constitutions that speak to this standard are Maryland (1776) and Delaware (1792). In each case, the upper houses were permitted to amend and design bills that "incidentally" raised revenue, not to expand their involvement in the taxing power of the legislature, but to prevent the lower houses from mixing non-revenue raising measures into revenue raising bills thereby dishonestly circumventing the upper-house's input. Furthermore, in Maryland's case as the only pre-ratification constitution to mention the concept of incidental revenue, it is arguable that the Senate was not even permitted to originally amend a money bill.

Eleven years after independence, the delegates to the Constitutional Convention in Philadelphia had a wealth of State experiences in Origination Clause codification and legislative implementation to guide the national debate. (See Table 1 below for summary of State constitutions)

The experiences of the Colonies and early States under royal charters, statutes, State constitutions and various organic laws forms the body of common law explaining the context of our current constitutional system of law. The preceding examination of that legal tradition with respect to the origination principle on the American continent is meant to add to our interpretive understanding of both the Clause in our current Constitution and the basis upon which it rests.

The principle of origination of taxing measures only through popular, locally representative assemblies was well established in the Americas, and widely codified. Its contravention served as the primary cause for revolution against the government of England.

The Origination Clause

Table 1: Origination Requirements and Early State Constitutions

	Constitution	Bicameral Legislature	Origination in Lower	Upper House may amend	Revenue Raising Definition	Senate Disposition
Connecticut	NA/Charter					
Delaware	1776	Yes	Yes	Yes	1792 addition	
Georgia*	1777	No				
Maryland	1776	Yes	Yes	Yes	Yes (strict)	Electoral College
Massachusetts	1780	Yes	Yes	Yes		
New Hampshire	1776	Yes	Yes	Yes (1784)		
New York	1777	Yes	No (explicit in 1821)			Indirect elections
North Carolina	1776	Yes	No			Annual elections
Pennsylvania	1776	No (Yes in 1790)	Yes (1790)	Yes (1790)		
Rhode Island	NA/Charter					
South Carolina	1776	Yes	Yes	No (Yes in 1790)		
Vermont**	1777	No				
Virginia	1776	Yes	Yes	No		

*upper house added in 1789, origination restraint in 1798

**upper house added in 1863, origination restraint and amending power added in 1863

IV. PHILADELPHIA 1787

In the interest of conserving space, we provide only a brief executive summary of the developments respecting the Origination Clause within the Constitutional Convention of 1787. For a more detailed chronological summary of the Origination Clause’s evolution in the Convention, the reader may view Krotoszynski’s 2005 article.⁶⁴ Our reading of the Convention’s journal and Madison’s notes diverge very little from Korotoszynski’s account of the development of the Clause.

⁶⁴ Krotoszynski, *supra* note 9, at 250-58.

When the Constitutional Convention opened on May 25th 1787, the fundamental topic of disagreement between the delegates that threatened progress towards amending the Articles of Confederation was over the nature of representation in the legislative branch. The smaller States were threatened by the Virginia plan's proposal of 29 May to proportion representation in the legislative branch according to population. George Read from Delaware threatened to "retire from the Convention" on the same day if the legislative principle of equal State representation under the Articles of Confederation was threatened. Charles Pinckney from South Carolina and Gouverneur Morris from Pennsylvania questioned whether altering the fundamental structure of the governing system under the Articles was even within the Congressional mandate for the Convention. The question was postponed in order to prevent "so early a proof of discord in the Convention as a secession of a State."⁶⁵

On 11 June, Roger Sherman of Connecticut opened by proposing the now famous Great Compromise providing for proportional representation in the House and equal representation in the Senate. Sherman cited as his example that "The House of Lords in England . . . had certain particular rights under the Constitution."⁶⁶ The issue of taxation and representation according to property contribution immediately took the debate. Benjamin Franklin's arguments were read aloud to the convention by his fellow Pennsylvania delegate, James Wilson: "The greater States Sir are naturally as unwilling to have their property in the disposition of the smaller, as the smaller are to have theirs in the disposition of the greater."⁶⁷ When the vote was put to allow equality of suffrage in the Senate, it initially failed (5-6) with the larger States generally voting against it. The smaller States had narrowly lost their first bid for equal representation.

On 13 June Elbridge Gerry of Massachusetts first moved to "restrain the Senatorial branch from originating money bills. The other branch was more immediately the representatives of the people, and it was a maxim that the people ought to hold the purse-strings. If the Senate should be allowed to originate such bills, they would repeat the experiment, till chance should furnish a sett of representatives in the other branch who will fall into their snares."⁶⁸ Pierce Butler from South Carolina disagreed in that there was no reason to mimic the tradition in the House of Lords and that it would lead to the "practice of tacking other clauses to money bills." Madison likewise disagreed arguing that the "the Senate would be the representatives of the people as well as the 1st branch." However, this was under the assumption that representation in the Senate was to be proportioned to population. South Carolina's delegate interrupted the debate on the wisdom of an origination restriction by pointing out

⁶⁵ MADISON, *supra* note 11, at 35-38 (May 29, 1787).

⁶⁶ *Id.* at 98.

⁶⁷ *Id.* at 101.

⁶⁸ *Id.* at 113.

The Origination Clause

that “the question [was] premature. If the Senate should be formed on the same proportional representation as it stands at present, they should have equal power, otherwise if a different principle should be introduced.”⁶⁹

In the face of the standoff, Benjamin Franklin interrupted with the first theoretical justification of the Origination Clause in his proposed compromise between the two groups:

The diversity of opinions turns on two points. If a proportional representation takes place, the small States contend that their liberties will be in danger. If an equality of votes is to be put in its place, the large States say their money will be in danger. When a broad table is to be made, and the edges of planks do not fit, the artist takes a little from both, and makes a good joint. In like manner here both sides must part with some of their demands, in order that they may join in some accommodating proposition.⁷⁰

Franklin’s ensuing compromise stated that in exchange for the small States getting equal representation in the Senate, that chamber would be restricted “generally in all appropriations & dispositions of money to be drawn out of the General Treasury; and in all laws for supplying that Treasury, the Delegates of the several States shall have suffrage in proportion to the Sums which their respective States do actually contribute to the Treasury.”⁷¹

On July 2nd, the delegates met to vote on equality of representation in the Senate without reference to an Origination Clause or Franklin’s proposed compromise. The resolution failed (5-5 with Georgia divided). With progress at a full stop, the members voted (9-2) to form a committee to detail a draft compromise following General Pinckney’s argument that “He liked better the motion of Doctr. Franklin (which see Saturday June 30). Some compromise seemed to be necessary: the States being exactly divided on the question for an equality of votes in the 2d. branch. He proposed that a Committee . . . be appointed to devise & report some compromise.”⁷²

The compromise committee worked through the 4th of July and on 5 July the delegates met again in convention to see the two part proposal they had produced:

I. That in the 1st. branch of the Legislature each of the States now in the Union shall be allowed 1 member for every 40,000 . . . that all bills for raising or appropriating money, and for fixing the Salaries of the officers of the Govern. of the U. States shall originate in the 1st. branch of the Legislature, and shall not be altered or amended by the 2d. branch: and that no money shall be drawn from the public Treasury. but in pursuance

⁶⁹ *Id.* at 114.

⁷⁰ *Id.* at 226-27.

⁷¹ *Id.* at 227.

⁷² *Id.* at 232.

of appropriations to be originated in the 1st. branch II. That in the 2d. branch each State shall have an equal vote.⁷³

The larger States' representative on the committee had "assented conditionally" to an equality of votes in the Senate, given that the "smaller States have conceded as to the constitution of the first branch, and as to money bills."⁷⁴ It is of historical significance that the prime bargaining chips used on this ultimate issue of intransigence in the Convention was a strict Origination Clause. The committee adjourned to consider the proposal the next day.

The Origination Clause was taken up for debate with various opinions on the necessity and wisdom of such a clause. Strong cases were made against the logic of an origination restriction on the Senate by Gouverneur Morris and James Wilson. George Mason and Franklin defended the necessity of the clause. According to Madison's records, the view that "generally prevail[ed]" was George Mason's argument that:

The consideration which weighted with the Committee was that the 1st branch would be the immediate representatives of the people, the 2nd would not. Should the latter have the power of giving away the people's money, they might soon forget the source from whence they received it. We might soon have an Aristocracy.⁷⁵

At the end of the debate on July 6th, the first draft of the Origination Clause without Senate amending power was voted for in the affirmative (6-3 with Georgia, New York, and Massachusetts divided). The following day, on July 7th the vote to allow equality of representation in the Senate for the small States was finally passed (6-3 with Georgia and Massachusetts divided). It had taken a month of heated debate that threatened to dissolve the Convention and the Union between the time of the proposal of the Virginia Plan and the actual compromise mechanism proposed by Benjamin Franklin including the Origination Clause that made progress possible. Specifically, it took the adoption of a strict Origination Clause against the Senate to convince enough of the larger States to allow equal representation in the Senate. The Origination Clause was the "cornerstone of the accommodation."⁷⁶

On July 16th the whole of the compromise was reaffirmed (5-4) in its complete legislative context:

[P]rovided always that representation [in the lower house] ought to be proportioned according to direct taxation; and . . . that all bills for raising or appropriating money, and for fixing the salaries of officers of the Govt. of the U. S. shall originate in the first branch of the Legislature of the U. S.

⁷³ *Id.* at 237.

⁷⁴ *Id.* at 242.

⁷⁵ *Id.* at 250.

⁷⁶ *Id.* at 290 (Gerry).

The Origination Clause

and shall not be altered or amended in 2d. branch [and] that in the 2nd branch of the Legislature of the U.S. each State shall have an equal vote.⁷⁷

Despite significant resentment and protest by several of the larger States that State equality in the Senate had been conceded, the Convention was finally able to move on to substantive discussions on the rest of the Constitution on July 17th. On July 26th, the Convention adjourned in order to allow the Committee of Detail to prepare a first draft of the whole Constitution for debate and revision. On August 6th, the committee produced the first draft with identical Origination Clause language as that cited in the 16 July vote above.

On August 8th, with the ink still wet on the first draft of the Origination Clause, Charles Pinkney and Gouverneur Morris motioned for a vote to repeal the clause completely citing that the Senate was competent to originate revenue bills, and that the clause would be responsible for “clogging the Government.[sic]” The hasty motion at the end of the day’s deliberation’s passed (7-4) with several of the smaller States voting for the repeal. The Convention adjourned for the day. As soon as the Convention opened the following morning, several representatives rose to express “dissatisfaction” with the clauses removal as its absence was “endangering the success of the plan, and extremely objectionable in itself.”⁷⁸ The absence of the Origination Clause continued to be a sticking point with several of the delegates as debates continued.

On August 11th following Edmund Randolph’s instance, a vote to reconsider the Origination Clause was taken up and passed (9-1). On the 13th, Randolph proposed to reinstate an amended Origination Clause with a narrower definition of revenue raising bills and a limited amending power in the Senate. The proposal read:

Bills for raising money for the *purpose of revenue* or for appropriating the same shall originate in the House of Representatives and shall not be amended or altered in the Senate as to increase or diminish the sum to be raised, or change the mode of levying it, or the objects of its appropriation.⁷⁹

The purpose of this amended clause as evidenced in the ensuing debate in the Convention was to prevent all potential bills that might “incidentally raise revenue”⁸⁰ from being excluded from Senate origination. To do this the lengthy compounded phrase “Bills for raising money for the *purpose of revenue* or for appropriating the same” was inserted with emphasis in the original added on the words “*purpose of revenue.*” In Mason’s mind this would remove Madison’s objection that all federal powers might have “some relation to money.” This is significant as the Supreme

⁷⁷ *Id.* at 297-98.

⁷⁸ *Id.* at 414 (Randolph).

⁷⁹ *Id.* at 442.

⁸⁰ *Id.* at 443 (Mason).

Court and others⁸¹ have since borrowed (knowingly or not) Mason's phrase, "incidentally raise revenue," in the body of Court precedent as the judicial standard for defining what is and is not considered a "revenue raising bill" in the context of the ratified Origination Clause of the Constitution. The Court has thrown out many past Origination Clause challenges against Senate originated taxes where revenue incidentally occurred in the Senate's pursuit of some other legitimate and enumerated "legislative ends" apart from taxing. However, two observations might cause the Court to pause if the framer's intent in the design of our mixed legislature is at all their guiding principle: First, the compound clause with emphasis on "*purpose of revenue*" was not adopted as proposed in Randolph's amendment. Second, in the very same paragraph, Mason clarifies that:

The Senate did not represent the *people*, but the *States* in their political character. It was improper therefore that it should tax the people. . . . Again, the Senate is not like the H. of Representatives chosen frequently and obliged to return frequently among the people. They are chosen by the Sts for 6 years, will probably settle themselves at the seat of Government will pursue schemes for their aggrandizement – will be able by [wearying] out the H. of Rep. and taking advantage of their impatience at the close of a long Session, to extort measures for that purpose. . . . If the Senate can originate, they will in the recess of the Legislative Sessions, hatch their mischievous projects, for their own purposes, and have their money bills ready cut & dried, (to use a common phrase) for the meeting of the H. of Representatives. . . . the purse strings should be in the hands of the Representatives of the people.⁸²

Additionally, the proposal here to relax the initial Origination Clause by allowing Senate amendments on bills not for the sole purpose of raising revenue would avoid the practice of the lower house "tacking foreign matter to money bills."⁸³ The addition of the Senate's amending power here was meant to alleviate fears that an aggressive House of Representatives might abuse an absolute origination prerogative on money bills by forcing the Senate to accept or refuse non-monetary statutes without their normal ability to amend or originate them.⁸⁴ The interpretation that the amending power was added to ensure the Senate had "some" taxing powers misses the actual and opposite concern the framer's had that an aggressive House might usurp the Senate's legitimate power over non-tax related statutes by

⁸¹ Additionally, Judge Joseph Story in his examination of what constitutes a revenue raising bill adopts the same understanding from the debates: "And, indeed, the history of the origin of the power, already suggested, abundantly proves, that it has been confined to bills to levy taxes in the strict sense of the words, and has not been understood to extend to bills for other purposes, which may incidentally create revenue. (citing Elliot debates)"

2 STORY, *supra* note 6, §877 at 343.

⁸² MADISON, *supra* note 11, at 443.

⁸³ *Id.* at 443-44.

⁸⁴ See 2 STORY, *supra* note 6, §872, 339-40. STORY acknowledges the same intent behind the tradition in the British Parliament.

proactively attaching non-taxing measures to House originated tax bills with the nefarious intent that the Senate could not alter them. James Wilson expressed the same concern that the House “will insert other things in money bills, and by making them conditions of each other, destroy the deliberative liberty of the Senate.”⁸⁵

Randolph’s proposal met significant skepticism in the Convention from several delegates. James Wilson was still against State equality in the Senate and argued for a bicameral legislature in which both houses were proportionally representative of the national population and controlled the purse strings equally. He saw in any Origination Clause only “a source of perpetual contention where there was no mediator to decide.”⁸⁶ Madison was supportive of the amended clause but foresaw extraordinary ambiguity in the language of the phrase “Bills for raising money for the *purpose of revenue*”:

The word revenue was ambiguous. In many acts, particularly in the regulations of trade, the object would be twofold. The raising of revenue would be one of them. How could it be determined which was the primary or predominant one; or whether it was necessary that revenue should be the sole object, in exclusion even of other incidental effects. . . . The words amend or alter, form an equal source of doubt & altercation. When an obnoxious paragraph shall be sent down from the Senate to the House of Reps it will be called an origination under the name of an amendment. The Senate may actually couch extraneous matter under that name. In these cases, the question will turn on the degree of connection between the matter & object of the bill and the alteration or amendment offered to it. Can there be a more fruitful source of dispute, or a kind of dispute more difficult to be settled?⁸⁷

Madison was for allowing the Senate amending power at least “to *diminish* the sum to be raised. Why should they [the Senate] be restrained from checking the extravagance of the other House.”⁸⁸

However, despite the theoretical challenges raised against the origination and amendment mechanism for revenue bills, one of the most persuasive arguments for retaining some sort of Origination Clause was purely pragmatic and popular. The Convention was mindful of the looming difficulties of ratification. Elbridge Gerry urged that the Convention retain the clause as it was:

[A] part of the plan that would be much scrutinized. Taxation & representation are strongly associated in the minds of the people, and they will not agree that any but their immediate representatives shall meddle with their

⁸⁵ MADISON, *supra* note 11, at 444.

⁸⁶ *Id.*

⁸⁷ *Id.* at 445-46.

⁸⁸ *Id.* at 445.

purses. In short the acceptance of the plan will inevitably fail, if the Senate be not restrained from originating Money bills.⁸⁹

John Dickenson echoed the same sentiment:

[A]ll the prejudices of the people would be offended by refusing this exclusive privilege to the H. of Reprers. . . Eight States have inserted in their Constitutions the exclusive right of originating money bills in favor of the popular branch of the Legislature. Most of them however allowed the other branch to amend. This he thought would be proper for U.S. to do.⁹⁰

Randolph stated a similar popular concern:

When the people behold in the Senate, the countenance of an aristocracy; and in the president, the form at least of a little monarch, will not their alarms be sufficiently raised without taking from their immediate representatives, a right which has been so long appropriated to them.--The Executive will have more influence over the Senate, than over the H. of Reprers--Allow the Senate to originate in this case, & that influence will be sure to mix itself in their deliberations & plans.⁹¹

On August 15th a new amended version with clearer Senate amendment prerogative was proposed by Caleb Strong of Massachusetts which read:

Each House shall possess the right of originating all bills, except bills for raising money for the purposes of revenue, or for appropriating the same and for fixing the salaries of the officers of the Govt. which shall originate in the House of Representatives; but the Senate may propose or concur with amendments as in other cases.⁹²

However, the Convention decided to postpone the issue until the specific powers of the Senate had been decided.

On September 5th, the Committee of 11 assigned to submit revised proposals for all postponed issues put forth the following Origination Clause language before the Convention: "All bills for raising revenue shall originate in the House of Representatives, and shall be subject to alteration and amendment by the Senate."⁹³ The language was taken up again on 8 September during the last day of substantive deliberation before the committee of style drafted the Constitution. Madison recorded no debate over the issue that day with the exception of the proposal to replace the phrase "and shall be subject to alteration and amendment by the Senate" with the language from Massachusetts's State constitution: "but the Senate may propose or concur with amendments as in other bills."⁹⁴ In both Madison's records and the Convention's journal, the only recorded vote was "On the

⁸⁹ *Id.* at 445.

⁹⁰ *Id.* at 447-48.

⁹¹ *Id.* at 448.

⁹² *Id.* at 461.

⁹³ *Id.* at 580.

⁹⁴ *Id.* at 607.

question of the first part of the clause – ‘All bills for raising revenue shall originate in the House of Representatives’.”⁹⁵ The vote on this first clause passed (9-2) with only Maryland and Delaware voting against it. The vote on the entire clause with the amending power was never officially recorded, however, a footnote in Madison’s records adds that “This was a conciliatory vote, the effect of the compromise formerly alluded to.” Regardless, the final language drafted by the committee of style for the delegates’ signatures included the amending power of the Senate.

V. THE MEANING OF WORDS

A. ORIGINAL PUBLIC UNDERSTANDING AND THE CHANGING JUDICIAL UNDERSTANDING:

While the debates within the Constitutional Convention are revealing of theory underlying the Origination Clause, of no less importance to the clause’s legal interpretation is the understanding of those who ratified it. The Convention was, after all, a meeting of delegates authorized only to propose amendments to remedy the inadequacies of the Articles of Confederation, and its proceedings were cloaked in secrecy from the general public for many years afterward. To ascertain the meaning and intent of the words of the ultimately-ratified Article 1, §7, we review what the words themselves meant to the public at the time, and the debate over its adoption in various public newspapers and proceedings during the period of ratification. It turns out that the judicial understanding has changed considerably overtime, and thus has often not matched the original public meaning of the clause.

The modern judicial interpretation of the words “revenue” and “originate” in the Origination Clause is controversial, and worth examination. The language stipulating the nature of permissible Senate amendments in the clause - “as on other bills” - warrants some examination as well.

i. “Revenue Raising”:

Consulting various period dictionaries for the definition of “revenue” from the late 18th through the early 19th century, one finds relatively uncontroversial meanings when compared with today’s connotation. In 1773 and 1799, the word “revenue” was defined in Samuel Johnson’s dictionary as “Income; annual profits received from lands or other funds.”⁹⁶ Although federal “revenue” may be thought to encompass something more than just tax revenue, the Origination Clause was about raising tax revenue, as Elbridge Gerry explained in a letter published in 1788 in which he protested against the Senate being able to amend revenue bills:

⁹⁵ *Id.* at 607.

⁹⁶ SAMUEL JOHNSON, A DICTIONARY OF THE ENGLISH LANGUAGE: in which the words are deduced from their originals, and illustrated in their different significations by examples from the best writers: to which are prefixed, a history of the language, and an English grammar (4th ed.1773); (8th ed. 1799).

[A] new provision now in the Constitution was substituted, whereby the Senate have a right to propose amendments to revenue bills & the provision reported by the committee was effectually destroyed. It was conceived by the committee to be highly unreasonable & unjust, that a small State which would contribute but one sixty fifth part of any tax should nevertheless have an equal right with a large state, which would contribute eight or ten sixty fifths of the same tax, to take money from the pockets of the latter, more especially as it was intended, that the powers of the new legislature should extend to internal taxation....⁹⁷

American usage of the word “revenue” in that era is also exemplified by the discussion in Federalist #12, written in 1787 by Alexander Hamilton:

In so opulent a nation as that of Britain, where direct taxes from superior wealth must be much more tolerable, and, from the vigor of the government, much more practicable, than in America, far the greatest part of the national *revenue* is derived from taxes of the indirect kind, from imposts, and from excises. Duties on imported articles form a large branch of this latter description. In America, it is evident that we must a long time depend for the means of *revenue* chiefly on such duties.⁹⁸

Hamilton’s linkage of “revenue” with taxes (including both direct and indirect taxes) continued to be reflected in American usage. For example, Webster’s American dictionary defined it in 1828:

In modern usage, income is applied more generally to the rents and profits of individuals, and revenue to those of the state. In the latter case, revenue is: 2. The annual produce of taxes, excise, customs, duties, rents, &c. which a nation or state collects and receives into the treasury for public use.

Further, the same 1828 American dictionary explains: “Government raises money by taxes, excise, and impost.” The combined words “revenue raising” were widely construed that way. Considering Hamilton’s and Webster’s use of the word “revenue,” it should be no surprise that the public would have understood revenue bills as those that tax in all the various forms of taxation. Additionally, it appears that it made little difference whether there was some intended legislative purpose or government program for the tax revenues. Franklin, the initial proponent of the origination mechanism in the Convention, himself confirms this in his memoirs when he repeatedly references the grant to fund Braddock’s Army in the French

⁹⁷ Elbridge Gerry, *Massachusetts Centinel* (Jan. 23, 1788) reprinted in THE DOCUMENTARY HISTORY OF THE RATIFICATION OF THE CONSTITUTION, Vol. 6, 1269 (hereinafter DHRC). Two hundred years later, Gerry’s state of Massachusetts was paying more federal taxes (\$21.7 billion) than the ten lowest states combined (\$21.4 billion). See Jim Luther, *Five Largest States Bear One-Third of Tax Burden, Group Says*, Associated Press (Apr. 29, 1986).

⁹⁸ THE FEDERALIST No. 12 (Alexander Hamilton).

The Origination Clause

and Indian War as a “bill for raising money.”⁹⁹ This understanding is also confirmed through the subsequent public debates over the Origination Clause where the term tax is treated interchangeably with revenue raising. For example, in 1788 an essay was published defending the proposed Origination Clause, stating that, “The people cannot be taxed, but, by the consent of their immediate representatives.”¹⁰⁰ (See Appendix A for an extended list of 20 examples of statements from the ratification debates regarding the Clause.)

The first edition of *Black’s Law Dictionary*, published in 1891, defines revenue this way: “As applied to the income of a government, this is a broad and general term, including all public moneys which the state collects and receives, from whatever source and in whatever manner. 22 Kan. 712.”¹⁰¹ The compound “Revenue laws” is next defined as, “Any law which provides for the assessment and collection of a tax to defray the expenses of the government is a revenue law. Such legislation is commonly referred to under the general term ‘revenue measures,’ and these measures include all the laws by which the government provides means for meeting its expenditures. 1 Woolw. 173.”¹⁰² In 1910, the second edition of *Black’s Law Dictionary* gives a definition for the compound term “Revenue Bills”: “These are the group of bills that impose the federal taxes. These bills originate in the House of Representatives.”¹⁰³ Likewise, one can look to the word’s adjacent use in the Constitution under article 1, §9 immediately following two taxing prohibitions: “No Preference shall be given by any Regulation of Commerce or Revenue to the Ports of one State over those of another.” For an even earlier understanding of what the American’s of the period considered inclusive in Revenue legislation, we could also return to the First Continental Congress’s explanation of the term’s scope in their 1774 declaration of rights and grievances:

[T]he British Parliament, claiming a power, of right, to bind the people of America by statutes in all cases whatsoever, hath in some acts expressly imposed taxes on them, and in others, under various pretenses, but in fact for the purpose of raising revenue, hath imposed rates and duties payable in these Colonies.

It is difficult to find any significant historical evidence that the early Americans considered the terms “revenue” and “revenue raising bills” to

⁹⁹ See, BENJAMIN FRANKLIN, MEMOIRS OF THE LIFE AND WRITING OF BENJAMIN FRANKLIN, 116, 121 (H. Colbern, 1818). Franklin earned significant fame in his early career in the Pennsylvania lower assembly opposing on Origination Clause grounds the attempts of the governor and proprietors to amend their money bills.

¹⁰⁰ Civic Rusticus, *Virginia Independent Chronicle*, 30 Jan 1788 reprinted in DHRC, *supra* note 97, V.8 at 335.

¹⁰¹ BLACK’S LAW DICTIONARY (A Dictionary of Law containing Definitions of the Terms and Phrases of American and English Jurisprudence, Ancient and Modern, 1st Ed.), 1040 (St. Paul, Minn. West Publishing Co. 1891).

¹⁰² BLACK’S LAW DICTIONARY 1040 (1st ed. 1891).

¹⁰³ BLACK’S LAW DICTIONARY (2d ed. 1910).

encompass only a narrow category of legislation. Their understanding seemed quite broad and inclusive of any act which might tax the people.

Despite this, the courts have since adopted multiple understanding of the term “revenue” and its compound. In the 19th century, the judicial interpretation seemed to coincided with the original understanding. For example, an 1875 federal Origination Clause case in New York stated:

Certain legislative measures are unmistakably bills for raising revenue. These impose taxes upon the people, either directly or indirectly, or lay duties, imposts or excises, for the use of the government, and give to the persons from whom the money is exacted no equivalent in return, unless in the enjoyment, in common with the rest of the citizens of the benefit of good government. It is this feature which characterizes bills for raising revenue. They draw money from the citizen; they give no direct equivalent in return. In respect to such bills it was reasonable that the immediate representatives of the taxpayers should alone have the power to originate them.¹⁰⁴

Likewise, in 1876, the U.S. Supreme Court held that a federal law authorizing post-offices to sell money orders did not raise revenue because the money was not obtained from levying taxes.¹⁰⁵ In 1887, the U.S. Supreme Court continued to closely link the word “revenue” with “taxation,” when construing the word “revenue” as used in federal statutes:

[T]he term “revenue law,” when used in connection with the jurisdiction of the courts of the United States, means a law imposing duties on imports or tonnage, or a law providing in terms for revenue; that is to say, a law which is directly traceable to the power granted to Congress by § 8, Art. I, of the Constitution, “to lay and collect taxes, duties, imposts, and excises.”¹⁰⁶

Supreme Court jurisprudence on this subject would soon take a sharp turn, 21 years later, so that some taxes would not qualify as revenue.

¹⁰⁴ *United States v. James*, 26 F. Cas. 577; Case No. 15464. (1875). This was an Origination Clause case, and the court held there was no violation of that clause, because charging money for postage stamps was not a tax.

¹⁰⁵ *United States v. Norton*, 91 U.S. 566 (1876). *Norton* mentioned the Origination Clause by way of analogy, but *Norton* was not an Origination Clause case. Rather, *Norton* involved interpretation of a statute written in 1804. See 2 Stat. 290, § 3 (Mar. 26, 1804). One might criticize *Norton* on the ground that all taxes are supposed to give taxpayers an equivalent in return, but the decision in *Norton* still seems well-justified since selling postal items to voluntary buyers rarely if ever amounts to taxation.

¹⁰⁶ *United States v. Hill*, 123 US 681 (1887). In 1844, Congress had allowed the Supreme Court to hear certain appeals regardless of the amount in dispute, but only if the appeal involved “the enforcement of the revenue laws of the United States.” Act of May 31, 1844, 5 Stat. 658. Of course, the United States adopted the Sixteenth Amendment in 1913, which created a further means of raising revenue.

ii. “Originate”:

In the 17th century, the word “original” meant “a beginning or fountain; An Original is also a first, authentick, or true draught of a writing.”¹⁰⁷ The term “draught (spelled “draft” nowadays) was defined in the period as “Delineation; sketch; outline.”¹⁰⁸

The term “origination” was defined in early America as simply “To bring into [or take] existence.”¹⁰⁹ In all the illustrative examples of the terms use from surveyed period dictionaries, there is usually some resemblance between the original and the resulting amended product. In this sense, it might have been normal to say that men originated from their ancestors. However, while perhaps technically true, it would be unconventional to say that men originated from water; so did every living organism, and it is of no use to describe water as the origination when doing so would confuse the audience by its lack of resemblance to the product.

Even today, it would be strange to say in plain English that a Senate-amended bill that is completely unrelated in substance to its House “shell bill”¹¹⁰ was in any sense “originated” by the House “shell bill.” The origination would be in formal numbering only, and such numbering has no constitutional significance, as it did not even exist in 1787.¹¹¹

The Framers and public were concerned about substantive taxes, not bill numbers/designators, and the Origination Clause attempts to alleviate

¹⁰⁷ THOMAS BLOUNT, GLOSSOGRAPHIA, or, A dictionary, interpreting the hard words of whatsoever language, now used in our refined English tongue: with etymologies, definitions, and historical observations on the same. . . (4th edition, London, 1674) (“Pedigree, or birth; a stock or kindred ; a beginning or fountain ; An Original is also a first, authentick, or true draught of a writing”). Of course, the outdated spelling of “draught” was equivalent to the modern spelling “draft.” The spelling “draught” was still used extensively in the colonial period in America as in Jefferson’s “original Rough draught” of the Declaration of Independence.

¹⁰⁸ SAMUEL JOHNSON, *supra* note 96, (1773): “Delineation; sketch; outline. ‘A good inclination is but the first rude draught of virtue; but the finishing strokes are from the will.’-South, ‘I have, in a short draught, given a view of our original ideas, from whence all the rest are derived.’-Locke”.

¹⁰⁹ NOAH WEBSTER, AN AMERICAN DICTIONARY OF THE ENGLISH LANGUAGE (1828).

¹¹⁰ A legislative vessel used by an amending chamber/body usually unrelated to the eventual product. Shell bills are amended by substitution by “strick[ing] all after the enacting clause”. We find little evidence that such a practice was endorsed in Parliamentary procedure on the American continent, and significant evidence that it was prohibited by rules and custom in the late 18th Century. Recently, the use of shell bills has become more common in the U.S. Congress and has generally been at least passively tolerated.

¹¹¹ The number referencing system did not even exist until 1817 in the House and 1847 in the Senate. The House adopted a sequential numbering system in which bills were numbered consecutively for an entire Congress in the 15th Congress (1817), and the Senate began using the same numbering system in the 30th Congress (1847). Prior to that time, the Senate numbering system provided that sequential numbering started anew at the beginning of each congressional session. *About Bills, Resolutions, and Laws*, available at www.lexisnexis.com/help/cu/Serial_Set/About_Bills.htm.

that substantive concern. It certainly would leave more than a few persons scratching their heads if Congress were to call a House bill the first draft of the resulting bill even after the Senate had substituted totally its own unrelated measure in place of the House bill.

The Court has said that it seeks to avoid impugning the character of members of a coequal branch, by questioning whether a formally enrolled bill, passed by each house, and signed by the president originated where Congress said it did.¹¹² More recently, however, federal courts have made clear that that standard is far from absolute.¹¹³

During the Virginia ratification debate in 1788, James Madison said that allowing Senate amendments would make it unnecessary for the Senate to “reject the bill altogether.” William Grayson replied that the Senate might claim power to reject the entire bill except one word, and add its own text instead, which Grayson said would be “the same, in effect, as that of originating.”¹¹⁴ Indeed, such an action by the Senate would be the same as originating, and Madison never suggested otherwise. On the contrary, Madison had taken the position that even changing a single paragraph of a bill could amount to an origination.¹¹⁵ Doubtless, the House and Senate have ultimate responsibility for determining what is and is not an origination, except in the most extreme cases, but the Senate has a strong motive to conclude that an amendment is not an origination because such conclusion increases the Senate’s power, and even the House has a motive to conclude that an amendment is not an origination (i.e. avoiding responsibility for taxes).

iii. Germaneness and the Phrase “as on other Bills”:

The Origination Clause specifies in the context of revenue-raising bills that, “the Senate may propose or concur with Amendments as on other Bills.” To ascertain the meaning of this phrase it is necessary to examine

¹¹² *Rainey v. United States*, 232 US 310 (1914) (“Having become an enrolled and duly authenticated Act of Congress, it is not for the Court to determine whether the amendment was or was not outside the purposes of the original bill”).

¹¹³ See *Munoz-Florez*, 495 U.S. at 389 n.2 (1990) (the Court “reserved the question whether there is judicial power after an act of Congress has been duly promulgated to inquire in which House it originated”).

¹¹⁴ Virginia Ratification Convention, Elliott 3:375-378 (June 14, 1788).

¹¹⁵ See Max Farrand, 2 THE RECORDS OF THE FEDERAL CONVENTION OF 1787, (Aug. 13, 1787), 273 (Yale U. Press 1937). Madison said:

When an obnoxious paragraph shall be sent down from the Senate to the House of Reps it will be called an origination under the name of an amendment. The Senate may actually couch extraneous matter under that name. In these cases, the question will turn on the degree of connection between the matter & object of the bill and the alteration or amendment offered to it. Can there be a more fruitful source of dispute, or a kind of dispute more difficult to be settled?

Id. If Madison had thought that the Senate could constitutionally introduce whatever extraneous matter it wanted, then he would not have expressed these concerns, nor would he have later extolled the power of the House on revenue matters (in Federalist #58).

the custom of the period. There were some norms of legislative procedure for upper house amendments that the ratifying public expected when they agreed to the Constitution.

Not only on revenue bills, but on all legislative acts, non-germane amendments were seen as an anathema. A substitute amendment is the most non-germane form of amendment conceivable. A “substitute amendment” is appropriately defined this way:

A motion, amendment, or entire bill introduced in place of the pending legislative business. Passage of a substitute amendment kills the original measure by supplanting it. The substitute may also be amended.¹¹⁶

The Senate was given amendment power primarily so that it could strip out non-germane provisions that the House might otherwise tack on to revenue bills. As Theophilus Parsons argued at the Massachusetts ratification convention, “had not the Senate this power, the representatives might tack any foreign matter to a money bill, and compel the Senate to concur, or lose the supplies.”¹¹⁷ Just as the Origination Clause inhibits tacking of foreign matter by the House, so too it places a limit on foreign matter tacked on by the Senate, by limiting the Senate to amendments rather than replacements, by forbidding the Senate to originate bills, and by requiring that Senate amendments be “as on other bills.”

The U.S. District Court for the District of Columbia, in the *Sissel v. HHS* case, recently concluded that any germaneness requirement for Senate amendments of House-originated revenue bills is a loose requirement at best.¹¹⁸ In her decision, Judge Howell relied on the fact that the Supreme Court approved of the Senate swapping a corporate excise tax for an inheritance tax in a revenue bill that the Senate had received from the House in *Flint v Stone Tracy* (1911). She wrote:

Although a corporate income tax is germane to an inheritance tax insofar as they are both taxes, the similarities end there. Hence, if the Supreme Court imposed a germaneness requirement in *Flint*, the most that it would require would be that both the original House bill and the Senate amendment be revenue-raising in nature.¹¹⁹

Actually, in *Flint*, the original House bill contained much more than the inheritance tax that was removed by the Senate. As the Court said in *Flint*: “the tariff bill, of which the section under consideration is a part,

¹¹⁶ RAMESH CHOPRA, ACADEMIC DICTIONARY OF POLITICAL SCIENCE, at 283 (2005). Chopra ironically lists this definition right next to the definition of the word “subversive” (“Tending to undermine, disrupt or supplant something already established. As in lawlessnessm...”).

¹¹⁷ See DEBATES, RESOLUTIONS AND OTHER PROCEEDINGS OF THE CONVENTION OF THE COMMONWEALTH OF MASSACHUSETTS, 126 (Boston, Oliver & Munroe, 1808).

¹¹⁸ *Sissel v. U. S. Dep’t of Health & Hum. Services*, No. 10-1263, D.D.C. June 28, 2013 (“*Flint* established a very loose conception of germaneness”).

¹¹⁹ *Id.* at 21.

originated in the House of Representatives and was there a general bill for the collection of revenue.”¹²⁰

The Payne Aldrich Tariff Act of 1909, into which the disputed corporate excise tax was written by the Senate, began in the House as a comprehensive tariff revision bill wholly designed by the House of Representatives.¹²¹ Along with the almost 900 tariff and excise schedules it affected, the House bill proposed an inheritance tax in the House’s original version after the House itself had already considered a corporate excise tax as an interchangeable substitute for the inheritance measure. The Senate (in cooperation with President Taft) thought it preferable to supplant the inheritance tax with a corporate excise tax. This was one item in a bill of hundreds of alterations to the U.S. tax structure. The Senate amendment was clearly germane to the subject matter *of the bill* that the House sent to the Senate, even if the removed clause was not germane to the inserted clause. The bill that entered the Senate was on the same subject and nearly identical to the bill that left the Senate after amendment. This was the context of the germaneness ruling in the *Flint* Court. No lengthy explanation of this point by the *Flint* Court was necessary for anyone familiar with the Payne Aldrich Tariff Act.

If there were no germaneness requirement, then the Origination Clause would be wholly superfluous, and furthermore the word “amend” in the Clause certainly does not mean “replace” in any dictionary of plain English. The nature of the amendment performed by the Senate in the 1909 tax bill was infinitesimal compared to that undertaken in the legislative history of the Affordable Care Act which the District Court defended in *Sissel v. HHS*. To create the Affordable Care Act, the Senate replaced a

¹²⁰ *Flint v. Stone-Tracy*, 220 U.S. 107, 143 (1911).

¹²¹ See Marjorie Kornhauser, *Corporate Regulation and the Origins of the Corporate Income Tax*, 66 IND. L.J. 63, 82, 93 (1990): “on March 4, 1909, Taft called for a special session of the 61st Congress, to convene March 15th, to deal with tariff reform and with the revenue need reported by the Secretary of the Treasury.” If tariffs could not provide enough revenue, he suggested that an inheritance tax should be enacted.” See also *id.* at 96: “Two versions were considered: one taxing dividends, the other taxing net earnings. As to either version, some saw a corporate tax as double tax, either in the sense that corporations already paid state taxes, or in the sense that a holding company would pay a tax when it distributed dividends to its shareholders on earnings already taxed when its operating companies distributed dividends to it so Representative Sereno Payne of New York told the House on March 23rd that the Committee had rejected a tax on the net earnings of corporations because many corporations were in a precarious financial condition (due to the Panic of 1907), states already taxed corporations, and the bill would not raise enough revenue. Congressman Longworth, speaking in July after the introduction of the corporate excise bill, stated that the Committee had rejected the proposal because it did not think the revenue was needed and also because the Committee had decided already to propose an inheritance tax, as suggested by Taft at his inauguration. At any rate the revenue bill that the House sent to the Senate for consideration consisted of the tariff provisions plus an inheritance tax but not a corporate tax.” See also Steven A. Bank, *Entity Theory as Myth in the Origins of the Corporate Income Tax*, 43 WILLIAM & MARY L. REV. 447 (2001).

714-word bipartisan (i.e. adopted by a vote of 416-0) House bill providing tax relief for veterans who were first-time homeowners,¹²² with a substituted 380,000-word bill including 17 new “Revenue Provisions” in its “Title X” section estimated to generate \$437.8 billion in net total revenue between 2010 and 2019.^{123 124 125} The Senate amendments in *Flint* and *Sissel* stand at opposite extreme ends of the spectrum. The Court in *Flint* only ruled on a case at one end of the spectrum in the 1911 case, and the other end of the spectrum is a matter of first impression.

The House of Representatives has not always enthusiastically defended its prerogatives under the Origination Clause, in part because avoiding responsibility for taxes is common behavior for members of Congress who must face the electorate every two years. Professor Kysar is thus technically correct in claiming that:

[T]he House gradually abandoned its restrictive view of the Senate’s amendment power. In fact, in 1909, no member of the House challenged the Senate’s conversion of the House tariff bill into a new tax on corporate income, at issue in *Flint*.¹²⁶

In addition to the political desire to avoid responsibility for taxation, the House’s 1909 behavior may also be partly explicable by the fact that the House had already considered the corporate excise tax as an alternative to the inheritance tax, and so the Senate amendment was not quite as foreign as if the House had never considered the idea.¹²⁷

In arguing against a germaneness standard, Kysar offers several legislative considerations for why the Court ought not to enforce a germaneness standard. Most of those considerations were already generously considered by the Framers and the public before giving the lower House the exclusive privilege of originating revenue measures while allowing the Senate to amend as on other bills. One particular consideration mentioned by Kysar is especially perplexing: the concern that “[a] strict germaneness requirement might prevent the Senate from excising an unrelated rider or otherwise threatening to retaliate against the House.”¹²⁸ We do not understand how deletion of an unrelated rider could ever be non-germane; if it is “un-

¹²² Service Members Home Ownership Tax Act of 2009.

¹²³ See John Cannan, *A Legislative History of the Affordable Care Act: How Legislative Procedure Shapes Legislative History*, 1 LAW LIBRARY JOURNAL, 105:2 (2013). Cannan’s article chronicles the unconventional legislative history of the Affordable Care Act, and the challenges such modern legislative procedures pose for researchers of the law.

¹²⁴ See Thomas.gov for legislative history of H.R. 3590 and its Senate Amendment 2786. Specifically §§ 9001-9017. Senate Amendment 2786 to H.R.3590 contained the vast majority of the substance of what would become the Affordable Care Act to include the majority of the bill’s revenue provisions.

¹²⁵ See Joint Committee on Taxation, *Estimated Revenue Effects . . .* (Mar. 20, 2010) https://www.jct.gov/publications.html?func=download&id=3672&chk=3672&no_html=1

¹²⁶ Kysar, *supra* note 10, at 32.

¹²⁷ See Kornhauser, *supra* note 121, at 82 *et seq.*

¹²⁸ Kysar, *supra* note 10, at 27-28.

related” then its deletion cannot possibly broaden the scope of the remaining material in the bill. Moreover, no one has argued that the judiciary should be free to address any but the most egregiously non-germane Senate amendments, leaving the remainder of germaneness decisions with Congress.

In determining the meaning of the phrase “as on other bills” in the context of the Origination Clause, a very useful reference would be the parliamentary procedures for amending and substituting bills during that era. In 1781, the Continental Congress passed this measure:

No new motion or proposition shall be admitted *under color of amendment as a substitute* for a question or proposition under debate until it is postponed or disagreed to.¹²⁹

This rule remained in effect in 1787 and 1788, and of course everyone understood at that time that the new U.S. Senate would be the successor body to the Continental Congress, representing states instead of population. Obviously, this rule of the Continental Congress would not allow erasure of a very popular bill to make room for an entirely different bill.¹³⁰

The first House of Representatives adopted the same rule as the Continental Congress in 1789. Though the Senate did not adopt that rule of its predecessor body, the House more than the Senate has responsibility for defending the House’s prerogatives under the Origination Clause, and the House has sometimes done so via its germaneness rules. The language for this House rule remained until its slight alteration in 1822: “No motion or proposition on a subject different from that under consideration shall be admitted under color of amendment.” According to the parliamentary precedents of the House of Representatives, “When therefore, it is objected that a proposed amendment is not in order because it is not germane, the meaning of the objection is simply that it (the proposed amendment) is a motion or proposition on a subject different from that under consideration.”¹³¹ In the 1780s (when the phrase “as on other bills” was written and publicly debated), the parliamentary custom in the national legislative body that preceded our bicameral legislature was clearly against the practice of gutting legislation to switch over to an entirely new text even though the gutted legislation has not been postponed or disagreed to.

In 1880, a point of order was made against an amendment to a bill being considered in the House on these grounds:

¹²⁹ ASHER CROSBY HINDS, PARLIAMENTARY PRECEDENTS OF THE HOUSE OF REPRESENTATIVES OF THE UNITED STATES, §1072 (U.S. Gov’t. Printing Office, 1899) (emphasis added).

¹³⁰ The House adopted the Service Members Home Ownership Tax Act of 2009 by a vote of 416-0, before it was gutted in the Senate. In the end, “the tax credit extension” for service members passed by using another failed bill. See Cannan, *supra* note 123, at 153, n.179 and accompanying text.

¹³¹ HINDS, *supra* note 129, at 568.

The Origination Clause

First, that it is not germane to the subject matter of the bill under consideration; and secondly, that it is in substance the same as a bill heretofore reported by the Committee on Printing and now pending before the House.¹³²

The Chair sustained the point of order, ruling that:

[E]ver since the 4th of March 1789, this House has had a rule which changed the common [British custom] parliamentary law in this respect, at least as to substitutes, and ever since 1822 as to amendments in any form. . . . after the bill has been reported to the House no different subject can be introduced into it by amendment, whether as a substitute or otherwise. . . . Since the adoption of the rule . . . in every instance where an amendment proposed to introduce an entirely new subject it has been excluded.¹³³

Likewise, on January 14th, 1898 a nearly identical parliamentary point of order was made against a substituted amendment and was sustained on the grounds that the subject was not germane to the bill under consideration.¹³⁴ In 1911, the House provided to its rules that,

[N]o amendment shall be in order to any bill affecting revenue which is not germane to the subject matter in the bill, nor shall any amendment to any item of such bill be in order which does not directly relate to the item to which the amendment is proposed.¹³⁵

The House has held non-germane countless substitute amendments which are too numerous to list here, and they have been amendments to Senate as well as House bills.¹³⁶

Of course, neither the House nor the Senate existed at the time of the writing or ratification of the Constitution, so their (since) adopted customs are of limited value for understanding the phrase “as on other bills.” That is especially true of the Senate, and not just because House prerogatives are at issue here; the Senate’s rules have been at odds with the American custom of the day as the Senate became “almost if not quite the only parliamentary body in this country adhering in any degree to the English belief that an amendment need not be germane.”¹³⁷

In matters of the purse, the Constitution indicates that the Senate amending power on revenue bills is limited “as on other bills,” meaning bills and amendment procedures understood by the Framers and public. Professor Kysar disagrees:

¹³² *Id.* at 568.

¹³³ *Id.*

¹³⁴ *Id.* at 569.

¹³⁵ LUCE, *supra* note 20, at 429.

¹³⁶ See HINDS, *supra* note 129, index under “Germane amendments: Decisions discussing at length the quality of germaneness in amendments Amendments must be germane . . . See Amendments... It is not in order to move to recommit as bill . . . which is not germane.”

¹³⁷ LUCE, *supra* note 20, at 429.

The Senate's power to amend has traditionally been quite broad. . . . The Senate has never had a rule against non-germane amendments and thus early congressional practice and American understanding of parliamentary practice leaves room for such freedom. Since the Senate possesses the power to attach non-germane amendments to non-revenue bills, the Constitution thus appears to prescribe its power to do so in the context of revenue bill.¹³⁸

However, the Senate only possesses the power to attach non-germane amendments to non-revenue bills because it gave that power to itself after the Constitution was ratified. That decidedly non-American tradition which the Senate unilaterally adopted was at odds with the common legislative requirement of the time that amendments be germane. Of course, the Senate is entitled to make its own rules, but there are constitutional limits. No one disputes, for example, that the Senate cannot take a House-originated bill unrelated to revenue, and convert it by amendment into a revenue-raising bill:

In 1864 when the House questioned the right of the Senate to provide a tax on incomes by amendment to a non-revenue bill, the Senate withdrew the amendment. In 1878 the House returned to the Senate a House bill about postroads to which the Senate had added revenue amendments, the House vote being 169-68. Speaking more emphatically with a unanimous vote, the House in 1905 sent back a bill relating to the taxation of bonds issued to aid isthmian canal construction. The Senate had stricken out all of the House bill after the enacting clause, and inserted somewhat similar provisions. A conference committee had restored the substance of the original House bill, but used the Senate language. Nevertheless the House insisted strictly on its prerogative.¹³⁹

A Senate amendment gutting a House revenue bill should be no more immune from constitutional scrutiny than a Senate amendment converting a non-revenue bill into a revenue bill. Both transgress the Origination Clause.

At least since Jefferson's Manual of 1801 was largely adopted by the Senate in 1828, the Senate has repeatedly rejected points of order challenging non-germane amendments to non-revenue bills. A civil rights bill was introduced in 1872 via a substitute amendment, and controversy ensued even though no revenue was involved:

Mr. SUMNER. . . I propose to move to strike out all after the enacting clause and insert what is generally known as the civil rights bill. . . I shall take the form of the bill which is now pending in the other house, which in substance and almost precisely in language is that on which the Senate acted. There are one or two verbal changes, but not important in principle or in any way affecting any principle of the bill. . . The VICE PRESIDENT. The Chair may say, in reply to the suggestion of the Senator from Connecticut [who had objected on grounds of the non-germaneness

¹³⁸ Kysar, *supra* note 10, at 29.

¹³⁹ LUCE, *supra* note 206, at 418.

The Origination Clause

of the substitute amendment], by which he enforces the point of order, that constitutional law and parliamentary law are often quite different. . . . but the Chair decides this question solely upon the parliamentary law applicable in this body. Now the Chair desires to add to this that by the parliamentary law as practiced in the House of Representatives, which is the parliamentary law as generally understood by Legislatures and parliamentary bodies in the United States, this amendment would be totally out of order.¹⁴⁰

Several items in these proceedings are notable. First, the Senator proposing the substitute amendment (Sumner) here felt obliged to communicate that the bill was not wholly originated by himself and that it was “in principle” identical to a house bill under consideration. Second, the Chair admitted that the sole authority governing his decision dismissing the germaneness objection was the Senate’s own rules and not constitutional considerations. And, third, the Chair admits that the Senate’s unilaterally-adopted rule allowing non-germane amendments would be under “parliamentary law as generally understood by Legislatures and parliamentary bodies in the United States. . . totally out of order.”

That entire 1872 Senate debate was over a rule governing a non-revenue raising bill, and therefore is subject to the Constitution’s allowance in article 1, §5 that “[e]ach House may determine the Rules of its Proceedings.” However, that rulemaking power is not unlimited, particularly where the Constitution specifies otherwise. The phrase “as on other bills” in the Origination Clause is just such a limitation. In the same sense, the Senate could not write a rule specifying when it could adjourn, or whether it had to keep a “Journal of Proceedings.” The Senate cannot adopt or employ rules for amending revenue-raising bills against the constitutional requirement that the amendments must be “as on other bills.”

Professor Kysar’s claim that, “[t]he Senate has never had a rule against non-germane amendments”¹⁴¹ is not quite accurate regarding the written Senate rules.¹⁴² Nor is it accurate when we consider that individual Senators have often felt obliged by the Origination Clause to limit their

¹⁴⁰ U.S. Senate Proceedings, in Blair & Rives, 66 *The Congressional Globe*, May 8 1872, Part 4 at 3181-83.

¹⁴¹ Kysar, *supra* note 10, at 29.

¹⁴² For example, the standing rules of the Senate today specify that on appropriations bills and amendments to appropriation bills there is a germaneness requirement:

On a point of order made by any Senator, no amendment offered by any other Senator which proposes general legislation shall be received to any general appropriation bill, nor shall any amendment not germane or relevant to the subject matter contained in the bill be received; nor shall any amendment to any item or clause of such bill be received which does not directly relate thereto

U.S. Senate Committee on Rules & Administration, *Appropriations and amendments to general appropriations bills*, (available at

www.rules.senate.gov/public/index.cfm?p=RuleXVI). Likewise, the current Senate also has a post-cloture germaneness rule (“No dilatory motion, or dilatory amendment, or amendment not germane shall be in order”). See Rule XXII, U.S. Senate.

amendments to germane ones. Just to take one example, in 1879, Senator James Beck, Democrat of Kentucky, raised a point of order, saying: “the amendment seeks to originate a revenue bill bearing upon external taxation... and as it is proposed as an amendment to an internal-revenue bill it is not germane to the bill.”¹⁴³ Senator Beck’s point of order lost, on a vote of 22 to 16,¹⁴⁴ but the larger message is that many Senators have felt themselves bound by a constitutional germaneness rule, even if they were a minority and even if the written Senate rules did not reflect that constitutional rule.

Since 1879, the Senate has grown more accustomed to wholly disregarding House-originated bills and supplanting them with their own meaning. Towards the end of the 19th century, the standard began to change as “[l]ittle by little the Senate accustomed itself to almost ignoring what the House sent over in the way of a money measure, and the country came to expect that the Senate will do no more than take a House bill for the foundation of its own structure.”¹⁴⁵ However, the U.S. Senate’s recent parliamentary philosophy does not represent the typical American experience and understanding of upper house amending power. By the middle of the 20th century, the U.S. Senate was “almost if not quite the only parliamentary body in this country adhering in any degree to the English belief that an amendment need not be germane.”¹⁴⁶ This philosophy embodied in Jefferson’s Manual apparently derived from fear that presiding officers would be stifling (e.g. that they would exercise too much control over the content of bills). However, as the Congressman and scholar Robert Luce put it, “[Jefferson’s] fears were unfounded, for often with little difficulty and rarely with harm nearly all American presiding officers now apply special rules requiring amendments to be germane.”¹⁴⁷

Even the U.S. Senate has formally recognized a germaneness requirement for some types of bills. They have historically instituted rules specifying that:

No amendment which proposes general legislation shall be received to any general appropriation bill, nor shall any amendment not germane or relevant to the subject-matter contained in the bill be received; nor shall any amendment to any item or clause of such bill be received which does not directly relate thereto. . .¹⁴⁸

¹⁴³ 8 Cong. Rec. 1478 (1879).

¹⁴⁴ *Id.* at 1482.

¹⁴⁵ LUCE, *supra* note 20, at 417.

¹⁴⁶ *Id.* at 429.

¹⁴⁷ *Id.*

¹⁴⁸ Congressional Serial Set, “Precedents of the Senate” at 60 (GPO 1914). Today a similar germaneness standard exists for Senate amendments to appropriation bills:

On a point of order made by any Senator, no amendment offered by any other Senator which proposes general legislation shall be received to any general appropriation bill, nor shall any amendment not germane or relevant to the subject matter contained in the bill be received; nor shall any amendment to any item or

The Origination Clause

Of course, the Senate did not exist before the ratification of the Constitution, so analyzing how the newly-formed Senate construed its own limitations may be of little avail in understanding what those who ratified the Constitution meant and consented to when considering the words “as on other bills.” Recall that the Senate chair said in 1872 that his determinations were bound “solely upon the parliamentary law applicable in this body” and not by constitutional considerations. Additionally, the historical absence of House or judicial opposition to Senate usurpations does not give such usurpations any form of constitutional legitimacy. As one scholar argued in concluding his 1919 examination of the Clause:

Should individuals and firms be protected against taxes adopted in an unconstitutional manner? It is not sufficient for the Court to declare that it is powerless to interfere, since the House has, perhaps under the stress of circumstances or unwittingly, assented to the Senate’s abuse of its privilege. Neglect cannot fairly be considered as an admission that trespass is justified.¹⁴⁹

We have explored the post-ratification Senate’s unique traditions here primarily to dispel historical misconceptions that there was a complete absence of any germaneness standard. We have done so in disagreement with a recent claim voiced in an academic publication and relied on by district court judges that “[t]he Senate’s power to amend has traditionally been quite broad. . . . The Senate has never had a rule against non-germane amendments and thus early congressional practice and American understanding of parliamentary practice leaves room for such freedom.”¹⁵⁰ Our review of early congressional practice and American understanding of parliamentary practice contradicts this claim and its implication that such broad amendment discretion must therefore extend to revenue raising bills. Moreover, where we do review the Senate’s early customs and traditions with respect to that chamber’s conception of its role in the design of money bills, we find significant evidence that the Senate has since its formative days viewed its own role with respect to such legislation as extraordinarily limited by custom and constitutional design:

clause of such bill be received which does not directly relate thereto; nor shall any restriction on the expenditure of the funds appropriated which proposes a limitation not authorized by law be received if such restriction is to take effect or cease to be effective upon the happening of a contingency; and all questions of relevancy of amendments under this rule, when raised, shall be submitted to the Senate and be decided without debate; and any such amendment or restriction to a general appropriation bill may be laid on the table without prejudice to the bill.

U.S. Senate Committee on Rules & Administration, *Appropriations and Amendments to General Appropriations Bills* (available at www.rules.senate.gov/public/index.cfm?p=RuleXVI). Likewise, the current Senate also has a post-cloture germaneness rule (“No dilatory motion, or dilatory amendment, or amendment not germane shall be in order”). See Rule XXII, U.S. Senate.

¹⁴⁹ Sargent, *supra* note 7, at 351-52.

¹⁵⁰ Kysar *supra* note 10, at 29.

As Haynes notes, “the Senate was not five months old when it denied to itself the power to originate a bill imposing an increased duty of tonnage.” A committee chaired by Senator Butler was appointed on June 17, 1789 to work on a bill “to arrange and bring forward a system to regulate the trade and intercourse between the United States and the territory of other powers in North America and the West Indies.” The committee reported the following on August 5th: “That it will be expedient to pass a law for imposing an increased duty of tonnage . . . but such a law being of the nature of a revenue law, your committee conceive that the originating a bill for that purpose, is, by the constitution, exclusively placed in the House of Representatives.” The Senate approved this report.¹⁵¹

The Courts may certainly be justified in generally deferring to each House to “determine the Rules of its Proceedings,” and to generally defend their own constitutional prerogatives. Such discretion is warranted on non-revenue-raising measures, and even for revenue-raising measures where the non-germaneness is less than extremely obvious. But when amendment practices are applied by the Senate to grant itself the power to effectively originate taxing provisions, the Constitution limits this practice – as much as it limits the Senate in transgressing any other constitutional limitations. Justice Thurgood Marshall, citing Federalist 58, said as much the last time the Supreme Court ruled on an Origination Clause claim:

Provisions for the separation of powers within the Legislative Branch are thus not different in kind from provisions concerning relations between the branches; both sets of provisions safeguard liberty. . . . A law passed in violation of the Origination Clause would thus be no more immune from judicial scrutiny because it was passed by both Houses and signed by the President than would be a law passed in violation of the First Amendment.¹⁵²

B. FURTHER EVIDENCE FROM THE RATIFICATION DEBATES:

We review a variety of sources from the ratification period from newspaper editorials to debates in the various legislatures.

We find only one definitive example of anyone raising the prospect of a Senatorial substitute amendment on a revenue bill in the thousands of collected public documents in *The Documentary History of the Ratification of the Constitution*. During the debates in the Virginia Legislature, one member objected to his understanding of the Origination Clause:

Mr. Grayson objected to the power of the Senate to propose or concur with amendments to money bills. He looked upon the power of proposing amendments to be equal in principle to that of originating, and that they

¹⁵¹ DANIEL WIRLS & STEPHEN WIRLS, *THE INVENTION OF THE UNITED STATES SENATE, 188-89* (2004). Internal citations omitted. Quotations cited internally in the original as: Haynes, 1938 at 432 and *Senate Journal*, Aug. 5, 1789 respectively.

¹⁵² *United States v. Munoz-Flores*, 495 U.S. 385, 395; 397 (1990).

The Origination Clause

were in fact the same. As this was, in his opinion, a departure from that great principle which required that the immediate Representatives of the people only should interfere with money bills; he wished to know the reasons on which it was founded. . . . Mr. Grayson still considered the power of proposing amendments to be the same in effect, as that of originating. The Senate could strike out every word of the bill, except the word Whereas, or any other introductory word, and might substitute new words of their own.¹⁵³

Madison himself responded to Grayson's fear that "amendment" was equivalent to "origination" by assuring him in a somewhat dismissive fashion that,

The criticism made by the Honorable Member, is, that there is an ambiguity in the words, and that it is not clearly ascertained where the origination of money bills may take place. I suppose the first part of the clause is sufficiently expressed to exclude all doubts. . . . Virginia and South-Carolina, are, I think, the only States where this power is restrained [no Senate amendment's to revenue bills]. In Massachusetts, and other States, the power of proposing amendments is vested unquestionably in their Senates. No inconvenience has resulted from it.¹⁵⁴

It is not astonishing that this explicit apprehension of the Senate using shell bills was contemplated by an Anti-Federalist member of the Virginia legislature. Virginia was one of the few exceptions among the States at the time in not having any experience with Senate amendments to revenue bills. George Mason, another Virginia Anti-Federalist (who was an actual member of the Constitutional Convention in Philadelphia), did not go so far as Mr. Grayson in stating his famous case against the Origination Clause. In Mason's passionate caution against the various grants of power contained in the new Constitution he warned that,

The Senate have the Power of altering all Money-Bills, and of originating Appropriations of Money, & the Salaries of the Officers of their own Appointment in Conjunction with the President of the United States; altho' they are not the Representatives of the People, or amenable to them.¹⁵⁵

It seems from Mason's warnings here that he distinguished a more extensive appropriation power in the Senate than taxing power by distinguishing "originating" from "altering" in each case. This seems to be the strongest case against article 1, §7 that Mason could conceive of. In reflecting on Mason's less ambitious attack on article 1, §7, Madison privately wrote to George Washington about Mason's objections regarding "the paltry right of the Senate to propose alterations in money bills."¹⁵⁶ If there is

¹⁵³ Virginia Convention Debates, 14 June 1788 *reprinted in* DHRC, *supra* note 97, V.10 at 1268.

¹⁵⁴ *Id.* at 1268.

¹⁵⁵ George Mason's Objections to the Constitution of Government formed by the Convention, *reprinted in* DHRC, *supra* note 96, V.13 at 43.

¹⁵⁶ James Madison to George Washington, New York, 18 Oct. 1787 *reprinted in* DHRC, *supra* note 97, V.13 at 408 (also in V.8 at 76).

any doubt about how Madison presented the clause to the ratifying public, his famous reflections in Federalist 58 was in Madison's own description of the first half of the clause, "sufficiently expressed to exclude all doubts": "The House of Representatives cannot only refuse, but they alone can propose the supplies requisite for the support of the government."¹⁵⁷

This singular example of a member of the ratifying public contemplating a broad Senate tax origination power through substitute amendment if taken in a vacuum would be the strongest case that such a power was understood by the public. However, when weighed against the body of contrary statements, it appears as an anomaly wholly refuted. Given the vast number of references during the ratification period evincing a more limited understanding of the Senate's amending power, we will confine ourselves to documenting 20 examples in "Appendix A" without room for extended discussion of each.

While all of the examples contained in "Appendix A" of the public's understanding of the Clause may have slight variations of interpretation, none of them premeditate a Senate's wholesale construction of tax bills. This is astounding given the wide and creative variety of apprehensions voiced by opponents of the Constitution in the heated ratification debates. What was plainly understood by article 1, §7 was that the Senate would be constitutionally restrained from designing taxes by the first half of the clause, and the House could not get away with tacking foreign matters to money bills by the second half of the clause.

Such is the context of the clause "All Bills for raising Revenue shall originate in the House of Representatives; but the Senate may propose or concur with Amendments as on other Bills." It was the primary bargaining chip used to bridge the disagreement between the large and small States that threatened progress on the Constitution. By privileging the proportionally representative House of Representatives on taxing measures, the Connecticut Compromise garnered enough support from the larger States to concede equal State representation in the Senate.

The Senate's power to amend revenue raising bills was added not as a compromise to those seeking to empower the Senate on taxing measures but as a means to avoid a disingenuous House of Representatives that might force the Senate to accept or refuse non-revenue related measures tacked onto revenue raising bills. However, its most fundamental role was as a pragmatic addition that was seen as alleviating popular prejudices against "taxation without representation" that divisive Senate originated tax bills might instigate.

The totality of the historical evidence from the ratification period indicates that almost no one expected that the clause would empower the Senate to legitimately originate taxes by unconventional (and illegal at the time) parliamentary amendment maneuvers. Furthermore, its plain under-

¹⁵⁷ THE FEDERALIST NO. 58 (James Madison).

standing was that it limited the upper branch from designing measures exerting control over the purse of the nation.

The Framers were fully aware of the enforcement and interpretive difficulties inherent in the clause. The controversies surrounding what constituted a “[b]ill for raising Revenue” were considered in the Convention, as well as the various evasive maneuvers each house might take to avoid the Clause’s requirement in their faction’s or chamber’s interests, or in collusion with the executive branch. However, despite all of these considered difficulties, the Framers decided to restrain the origination of all revenue raising bills (without the emphasis of “for the purpose of revenue”) to the more popular and nationally representative chamber. They allowed the Senate amending power primarily to prevent the popular branch from abusing a strict origination privilege in the absence of Senate amending power.¹⁵⁸

VI: SUPREME COURT PRECEDENT

Article I, Section 7, Clause 1, provides that: “All Bills for raising Revenue shall originate in the House of Representatives; but the Senate may propose or concur with Amendments as on other Bills.” The preceding historical analysis indicates concerns for the institutional structure for revenue raising among the Framers. An analysis of judicial interpretation offers little in the way of interpretive clarity. The Supreme Court of the United States has only waded into Origination Clause matters on only a handful of times in its history.¹⁵⁹ Lower courts have discussed origination with more frequency, but more often than not have deferred to the legislative branch by claiming that the law in question was not a revenue raising bill

¹⁵⁸ We diverge from Krotoszynski’s understanding of Senatorial taxing power as intended by the framers. See p. 259 of his 2005 article, *supra* note 10: “The only question presented for consideration was whether the failure to apportion Senate seats based on population made the Senate sufficiently similar to the House of Lords to justify strict limits on the body’s ability to influence fiscal policies directly. Notwithstanding the objections offered by Gerry, Mason, and Randolph, the delegates concluded that the Senate’s manner of selection and apportionment did not require limiting its voice in matters of taxing and spending.” Gerry, Mason, and Randolph’s views on origination issues were the ones that “generally prevailed” in the Convention according to Madison himself. Additionally, apportionment was not the only reason why the Senate was deemed unfit to tax the people. Term lengths, indirect elections, and the ratio of representatives to constituents were equally if not more often cited.

¹⁵⁹ See case file from legallanguage.explorer.com. This is similar to a Shepard’s citation search in Lexis-Nexis. The Supreme Court has decided only six substantive Origination Clause cases: *Rainey v. United States*, 232 U.S. 310 (1914); *Flint v. Stone Tracy Co.*, 220 U.S. 107 (1911); *Millard v. Roberts*, 202 U.S. 429 (1906); *Twin City Bank v. Nebecker*, 167 U.S. 196 (1897); *United States v. Norton*, 91 U.S. 566 (1875); and *United States v. Munoz-Flores*, 495 U.S. 385 (1990). Several other Supreme Court cases mention the Origination Clause, but only in passing. See, e.g. *Skinner v. Mid-America Pipeline Co.*, 490 U.S. 212 (1989), *United States v. Sperry Corp.*, 493 U.S. 52 (1989), and *Nat’l Fed’n of Indep. Bus. v. Sebelius*, 567 U.S. ____ (2012) (Scalia, J., dissenting).

or that the Senate's amendment at issue was germane to the subject matter of the House-originated revenue raising bill.¹⁶⁰

What seems to capture the Supreme Court's attention is the question, "what is a revenue raising bill?" On a few occasions, the Court's decision hinges on defining what makes a bill a revenue raiser. When concern is raised in Congress about tripping the Origination Clause language, the House and Senate have dealt with the provision by using a process called "blue slipping" in the House and in the Senate by a "question . . . submitted directly to the Senate," to indicate that a violation of the Origination Clause has taken place.¹⁶¹ These internal norms should be considered in light of their relationship to constitutional triggers of the Origination Clause.

Historically, revenue raising bills (1) impose taxes upon the people – direct or indirect, (2) lay duties impost or excises for the use of the government, and (3) give the person from whom the money is extracted no equivalent in return unless commonly felt by the benefit of good government.¹⁶² In order for the Origination Clause to apply, the raising of money must be the bill's primary purpose rather than an incidental effect of the legislation, and the resulting funds must be for expenses or obligations of government generally rather than a single specific purpose. In the cases heard to date, the Supreme Court has narrowly interpreted "raising revenue" so that a statute that generates monies for a specified legislative function/program is not deemed to be a revenue raiser under the Origination Clause.¹⁶³ There must be "no purpose by the act, or by any of its provisions, to raise revenue to be applied in meeting the expenses and obligations of the government."¹⁶⁴

In *United States v. Norton*,¹⁶⁵ the Supreme Court decided on a Congressional act that created a postal money order system. An Act of Congress entitled "An Act to establish a postal money order system" did just that. Despite the fact that the revenue raised from the postal money order system went into the general Treasury, the Court focused on the intent of Congress to establish a money order system rather than the incidental effect of generating revenue. The litigation involved a clerk, Norton, who was employed in a New York money order office and who had been indicted for embezzlement. The primary focus of Norton's legal challenge was the statute of limitations. A violation of federal law attached a two year statute of limitations. However, if the statute had been determined to be a revenue raising provision, the statute of limitations for prosecution

¹⁶⁰ *Id.*

¹⁶¹ See James Saturno, *The Origination Clause of the U.S. Constitution: Interpretation and Enforcement*, Congressional Research Service, 9-10, Mar. 15, 2011 (<http://www.fas.org/sgp/crs/misc/RL31399.pdf>).

¹⁶² See *United States ex rel. Michels v. James*, 26 Fed Case, 577, 578 (C. C. N.Y. 1875).

¹⁶³ Krotoszynski, *supra* note 9, at 248.

¹⁶⁴ *Twin City Bank v. Nebecker*, 167 U.S. 196 (1897).

¹⁶⁵ 91 U.S. 566 (1875).

would have extended to five years. The Court ultimately ruled that the statute was not a revenue raising law. The Court in *Norton* discounted the revenue raising component to the Act, generating money that went to the general Treasury, instead focusing on Congress' goal of establishing a money-order system:

The Constitution of the United States, Art. I, Sec. 7, provides that "All bills for raising revenue shall originate in the House of Representatives." The construction of this limitation is practically well settled by the uniform action of Congress. According to that construction, it "has been confined to bills to levy taxes in the strict sense of the words, and has not been understood to extend to bills for other purposes which incidentally create revenue." Story on the Const., sec. 880.¹⁶⁶

In *Twin City v. Nebeker* (1897)¹⁶⁷ the Court focused on a provision of the National Banking Act of 1864. The goal of the legislation was to create a national currency. In doing so, the Act imposed certain taxes on bank notes in circulation. The Act at issue was passed on June 3, 1864. The bill began in the House of Representatives but the provision to include a tax was added by Senate amendment and was later agreed upon by the House. The Court held that a fee imposed on banks based on the average amount of notes in circulation was not a revenue raising bill, hence did not trigger the Origination Clause. The litigation focused on the National Banking Act, whose primary purpose was to establish a national currency. The Court reasoned that the primary goal was not to raise revenue and the resulting funds were wholly incidental to the Act's purpose and were not for use by the government generally. The opinion written by Justice John Marshall Harlan handily disposed of the problem of origination in a 9-0 decision. The Supreme Court ruled that, "the case is not one that requires either an extended examination of precedents or a full discussion as to the meaning of the words in the Constitution 'bills for raising revenue'."¹⁶⁸ Simply put, the tax imposed was incidental to the object of creating a national currency. The tax was a means for creating a currency system pursuant to Congress's Art. 1, §8 power to "To coin Money, [and] regulate the Value thereof," not for raising revenue for the government. The Court never reaches the merits of the case, whether or not the tax imposed by the Treasurer of the United States was unlawful: "An act of Congress providing for a national currency secured by a pledge of bonds of the United States, and which, in the furtherance of that object, and also to meet the expenses attending the execution of the act, imposed a tax on notes . . . is clearly not a revenue bill . . . The tax was a means for effectually accomplishing the great object of giving the people a currency. . . . There was no purpose by the act, or by any of its provisions, to raise revenue to be ap-

¹⁶⁶ *Id.* at 568-69.

¹⁶⁷ 167 U.S. 196 (1897).

¹⁶⁸ *Id.* at 202.

plied in meeting the expenses or obligations of the government."¹⁶⁹ Taken together, *Norton* and *Nebeker* focus more on the intent of Congress to carry out one of its enumerated powers over the money generated and its use. Showing deference to Congress' legislative action and purpose, *Norton* and *Nebeker* narrow the Origination Clause aperture, discounting the rich history of the Framers.

In *Millard v. Roberts*,¹⁷⁰ the Court upheld the constitutionality of property taxes that were imposed to fund a railroad terminal in the District of Columbia. The taxes, imposed to improve the rail system, were not levied to raise revenue but for the program put in place. Using the same logic as *Nebeker*, the Court ruled that the taxes were not for raising revenue but only for the stated purpose in the Act, hence those taxes did not raise questions under the Origination Clause. *Millard*, also a unanimous decision, relied heavily on the logic of *Nebeker*, as the Court quoted it extensively. *Millard* thus reaffirmed the understanding of the Origination Clause from *Norton* and *Nebeker*.¹⁷¹

The concept of "incidentally create[d] revenue" introduced in 1875 as *dictum* in *Norton* and then applied to the holding in *Nebeker* 22 years later deserves an expanded examination here. It is one of the most often cited and least scrutinized concepts in Origination Clause jurisprudence. It has had a profound impact upon the judicial interpretation of the Origination Clause because, where relied upon, it dramatically narrows the Court's understanding of what constitutes a revenue-raising bill to exclude "bills for other purposes which may incidentally create revenue."¹⁷²

However, the standing authority of the concept of "incidentally create[d] revenue" introduced judicially by *Norton* and *Nebeker* Court is not beyond question. In 1915, a federal court in New York struck down an Act of Congress as violative of the Origination Clause in *Hubbard v. Lowe*,¹⁷³ even though the tax at issue was not designed for raising general revenue, but rather was meant to discourage certain people from making certain cotton futures contracts. The *Hubbard* case was the only case in

¹⁶⁹ 167 U. S. 196, 202 (1897).

¹⁷⁰ 202 U.S. 429 (1906).

¹⁷¹ *Millard v. Roberts*, 202 U.S. 429 (1906). During the next decade, the Court acknowledged in two separate cases that the challenged tariff bills were revenue bills, but upheld the power of the Senate to amend them. See *Flint v. Stone Tracy Co.*, 220 U.S. 107 (1911); *Rainey v. United States*, 232 U.S. 310 (1914). Following *Millard* in 1906, the Court did not have occasion to follow the logic of *Nebeker* regarding incidental revenue until 1990, but that 1990 case did not involve any tax as had *Nebeker* and *Millard*. See *United States v. Munoz-Flores*, 495 U.S. 385 (1990) (monetary "special assessment" on persons convicted of a federal misdemeanor).

¹⁷² *Twin City v. Nebeker*, 167 U.S. 196, 203 (1897).

¹⁷³ *Hubbard v. Lowe*, 226 F. 135 (S.D.N.Y. 1915), *appeal dismissed mem.*, 242 U.S. 654 (1916). The law in question (the Cotton Futures Act) was reenacted following proper procedures on August 11, 1916. Solicitor General Davis therefore moved for dismissal of his appeal, and the Court obliged, calling the case "disposed of without consideration by the court." 242 U.S. 654 (1916).

U.S. history where a federal court invalidated a statute for violating the Origination Clause. The judge in *Hubbard* specifically dismissed the reasoning of the U.S. Supreme Court in *Nebeker*, saying that it (along with similar lower court decisions) required “a good deal of mental strain,”¹⁷⁴ and the judge instead cited a later U.S. Supreme Court decision for the proposition that the motive or purpose of Congress is not relevant to judges.¹⁷⁵ The court reasoned that taxes create revenue, so the Origination Clause applies, whatever the particular purposes and motives of Congress might be.

By the 1930s, the courts’ interpretation about incidental revenue-raising was still somewhat mixed, with “[s]ome hav[ing] excluded incidental revenue; some hav[ing] extended the constitutional provision to cover all revenue.”¹⁷⁶ In recent decades, courts and conventional legal opinion have relied heavily on *Nebeker*’s conception of incidental revenue, especially when taxation is not involved, and when Congress is incidentally raising revenue pursuant to non-tax powers.¹⁷⁷

But what was the authority for the Court’s adoption of the concept of “incidentally create[d] revenue” introduced in *Nebeker*? *Nebeker* and its progeny (including *Millard*) involved taxes that the Court has exempted from the word “revenue” in the Origination Clause. While such a narrowing of the word’s original meaning (See preceding examination of the early American understanding of the scope of “Revenue raising Bills”) may at first seem somewhat inexplicable, it appears that the Court in *Nebeker* was attempting to follow the relevant discussion in Justice Joseph Story’s 1833 Commentaries, which in turn may be viewed in the context of the Constitutional Convention of 1787, and the Maryland Constitution of 1776. This chain, however, has several weak links.

The concept of incidental taxation is specified nowhere in the Constitution, and it was both discussed and rejected at the 1787 Convention. The Court’s reliance on that concept in *Nebeker* and *Millard* can be traced back through Joseph Story’s writings on the subject, through three usages of the term in the Constitutional Convention, and possibly back to Maryland’s usage in its Origination Clause of 1776. Some might argue that there is a broad conceptual gulf between bills that intend solely to tax people, and bills that happen to tax people. Such a distinction between incidental revenue and revenue proper does not appear to be historically justified, especially if the revenue comes from “strict taxes” rather than other sources. To the populace paying the resulting taxes, the distinction seems

¹⁷⁴ *Hubbard*, 226 F. 135, 140.

¹⁷⁵ See *McCray v. United States*, 195 U.S. 27 (1904) (congressional motives are not relevant to judges); *United States v. O’Brien*, 391 U.S. 367, 383 (1968) (same principle).

¹⁷⁶ LUCE, *supra* note 20, at 406.

¹⁷⁷ See *United States v. Munoz-Flores*, 495 U.S. 385 (1990). *Munoz-Flores* involved a monetary “special assessment” on persons convicted of a federal misdemeanor, so no tax was involved, and moreover Congress was seeking the funds in support of its power to punish misdemeanors involving the immigration laws.

wholly irrelevant. Moreover, an exception for “incidental” taxes turns the Origination Clause into a “formal accounting” gimmick, because the Senate can always pair up taxing and spending provisions so as to avoid the Clause.¹⁷⁸

The opinion in *Nebeker* cited Story’s Commentaries at §880¹⁷⁹ which used the “incidental” revenue language while citing the 1787 Convention (“2 Elliot’s Debates, 283, 284”).¹⁸⁰ However, the Convention delegates were at that time rejecting – not accepting – a proposal by Edmund Randolph to eliminate the origination requirement for revenue-raising that was not “for the purpose of revenue.”¹⁸¹ Regarding that rejected proposal, Randolph, Mason and Madison all referred to the proposal as excluding “incidental” revenue-raising from the requirements of the Origination

¹⁷⁸ *Munoz-Flores* at 407-408 (Stevens, J., concurring).

¹⁷⁹ According to §877 of volume 2 of Story’s Commentaries from the original 1833 edition:

[T]he history and origin of the power, already suggested, abundantly proves, that it has been confined to bills to levy taxes in the strict sense of the words, and has not been understood to extend to bills for other purposes, which may *incidentally create revenue*. No one supposes that a bill to sell any of the public lands, or to sell public stock, is a bill to raise revenue, in the sense of the constitution. Much less would a bill be so deemed, which merely regulated the value of foreign or domestic coins, or authorized a discharge of insolvent debtors upon assignments of their estates to the United States, giving a priority of payment to the United States in cases of insolvency, although all of them might incidentally bring revenue into the treasury.

2 STORY, *supra* note 6, §877 (emphasis added). Notice that none of the four examples given here by Story (in the last two quoted sentences) involves any tax whatsoever.

¹⁸⁰ At 2 Elliot’s debates, 283, 284 we find no relevant material to the topic of incidental taxation. We assume Story was referencing other relevant material from August 1787, elsewhere in that same volume.

¹⁸¹ On July 26, 1787, the Convention adjourned to await a draft Constitution from the Committee of Detail, which was tasked with reflecting the agreements that had already been struck. On August 6, the draft arrived, including this: “All bills for raising or appropriating money, and for fixing the salaries of the officers of government, shall originate in the House of Representatives, and shall not be altered or amended by the Senate.” That draft was voted down, and the Convention adjourned to think it over. On August 11, Randolph proposed “a clause specifying that the bills in question should be for the purpose of Revenue, in order to repel ye. objection agst. the extent of the words ‘raising money,’ which might happen *incidentally*....” (emphasis added). Then on August 13, George Mason endorsed Randolph’s proposal: “This amendment removes all the objections urged agst. the section as it stood at first. By specifying *purposes of revenue*, it obviated the objection that the Section extended to all bills under which money might *incidentally arise*” (emphasis added). James Madison disagreed: “In many acts, particularly in the regulations of trade, the object would be twofold. The raising of revenue would be one of them. How could it be determined which was the primary or predominant one; or whether it was necessary that revenue shd: be the sole object, in exclusion even of other *incidental effects*” (emphasis added). Thus, the Randolph proposal (excluding incidental revenue by using the words “*purpose of revenue*”) was never included in the Constitution, and the August 6 language “bills for raising” was adopted.

Clause.¹⁸² Additionally, that proposal failed. That failed proposal was quite different from the one ultimately ratified in the Constitution (which does not specify that the scope of “Bills for raising Revenue” only includes those for the “*purpose of revenue*”). To the extent that Story may have been relying upon the 1787 deliberations for the notion that incidental taxes are exempt from the requirements of the Origination Clause, Story must have been mistaken, and the *Norton* and *Nebeker* Courts equally mistaken to follow suit.

Both Randolph and Mason were willing to exclude incidental revenue-raising from the requirements of the Origination Clause. Madison opposed them due to the impracticality of determining which congressional purposes were incidental and which were not. Therefore, the Origination Clause as it stands now does not attempt to distinguish between congressional purposes. As Judge Hough reasoned when he struck down the Cotton Futures Act in 1916: “It is immaterial what was the intent behind the statute; it is enough that the tax was laid, and the probability or desirability of collecting taxes is beside the issue.”¹⁸³

Even if the Framers’ debate in the Convention does not support the judicial concept of “incidentally create[d] revenue,” it is hypothetically possible that Randolph’s ill-fated proposal may have been inspired by Maryland’s then-unusual treatment of the subject in its State constitution (the Delaware Constitution of 1792 also employed the concept but post-dates the timeframe). Maryland’s 1776 constitution specified:

[T]hat no bill, imposing duties or customs, for the mere regulation of commerce, or inflicting fines for the reformation of morals, or to enforce the execution of the laws, by which an incidental revenue may arise, shall be accounted a money bill: but every bill, assessing, levying, or applying taxes or supplies, for the support of the government, or the current expenses of the state, or appropriating money in the treasury, shall be deemed a money bill.

If this was the basis of the (rejected) Randolph proposal, one would expect the Maryland delegates to have contributed extensively to this debate. Daniel Carroll (the cousin of Charles Carroll of Carrollton who drafted much of the Maryland Constitution alongside Samuel Chase), was present but only offered that the distinction would cause trouble: “The most ingenious men in Maryland are puzzled to define the case of money bills, or explain the Constitution on that point; tho’ it seemed to be worded with all possible plainness & precision. It is a source of continual difficulty & squabble between the two houses.”¹⁸⁴ At any rate, there is little evidence that the concept of “Bills for raising Revenue,” adopted by the 1787

¹⁸² *Id.*

¹⁸³ *Hubbard v. Lowe*, 226 F. 135, 137 (S.D.N.Y. 1915), *appeal dismissed mem.*, 242 U.S. 654 (1916).

¹⁸⁴ *Id.* at 449 (Statements offered by Daniel Carroll, one of Maryland’s delegates in Convention).

Convention and ratified by the public, actually imported Maryland's understanding and experience more than it did the other dozen States' respective understandings. It is much more reasonable to presume that the Origination Clause imported the understanding of all the other States which had no such limited conception of the scope of "Bills for raising Revenue," and made no mention of "incidentally create[d] revenue." Any influence that the Maryland provision may have had on the 1787 Convention ended with the Convention's rejection of the Randolph proposal. Even if we were to take Maryland's use of the concept of "incidentally create[d] revenue" as a source of continuing influence, it was limited to three specific categories of bills only, and it is not clear that the upper house in Maryland could even originally amend a money bill.

Even if the concept of "incidentally create[d] revenue" remains judicially relevant, the Court has never made clear what the term means. The legal scholar and judge must return to the authority cited by the *Nebeker* Court (Judge Joseph Story) to justify the concept's adoption and what he meant by it at that time. The term "incidental" was defined in the late 18th century as "[i]ncident; casual; happening by chance; not intended; not deliberate; not necessary to the chief purpose."¹⁸⁵ In Joseph Story's time it was defined as "1. Happening; coming without design; casual; accidental; as an incidental conversation; an incidental occurrence. 2. Not necessary to the chief purpose; occasional."¹⁸⁶ However, in the parlance of constitutional law, the word "incidental" often referred more specifically to implied powers that result from expressly enumerated powers. Thus, when Story (in disagreement with another scholar of his time, St. George Tucker)¹⁸⁷ was distinguishing "incidental" revenues from other revenues, Story's examples all involved no taxes whatsoever. Instead, Story pointed only to funding sources incidental to the non-tax powers of Congress: "selling public lands, or public stock," or bills "establishing the post-office, and the mint, and regulating the value of foreign coin." These examples speak solely to funding that is necessary and proper for executing specifically-enumerated powers other than the taxing power, which suggests that Story probably had no intention of exempting "incidental" taxes (in the strict sense of the word) from the coverage of the Origination Clause (i.e. he was misconstrued by the Court in *Nebeker* and ensuing opinions utilizing his

¹⁸⁵ SAMUEL JOHNSON, *supra* note 96, (1773).

¹⁸⁶ NOAH WEBSTER, *supra* note 109, (1828).

¹⁸⁷ STORY, *Commentaries*, §877 (citing ST. GEORGE TUCKER, *BLACKSTONE'S COMMENTARIES*, 261 (1803)). Judge Tucker wrote that "acts for establishing the post-office" should have originated in the House. We do not endorse Tucker's position, to the extent that he suggested that the Origination Clause could apply to revenue other than taxes.

concept of “incidentally create[d] revenue” to exempt tax measures from “raising Revenue” status and thus Origination Clause scrutiny).¹⁸⁸

Even if Story’s passage is given an interpretation that aggrandizes Senate power (which we believe unwarranted), still there is no support in that passage for the Senate to originate bills that balance the budget, or that extract revenue to pay for the general expenses of government, or that impose taxes unconnected to any enumerated power other than the taxing power. Likewise, even if we were to accept Justice Harlan’s 1897 decision in *Nebeker* (citing Story) as correct and still binding – despite the opinion in *Hubbard v. Lowe* deeming *Nebeker* to no longer be controlling – still Harlan was careful to note in *Nebekar* that, “[t]here was no purpose by the act or by any of its provisions to raise revenue to be applied in meeting the expenses or obligations of the Government.”¹⁸⁹

While the three cases *Norton*, *Nebeker*, and *Millard* indicate a narrowed Court view of the scope of the term “Bills for raising Revenue,” focusing on Congress’ intent when legislating, the ensuing Origination cases considered the procedural problem of House-Senate origination, amendments, and germaneness. In 1911 the Court decided *Flint v. Stone Tracy Company*.¹⁹⁰ Argued in 1910, reargued in 1911 and decided in the same year, the Court ruled 7-2 regarding a tax bill. The House of Representatives passed a comprehensive tax bill which included an inheritance tax. When the bill was taken up in the Senate, the upper house inserted a corporate tax in place of the inheritance tax. The bill was later passed as amended, however the corporations, Stone Tracy and fourteen others, claimed the tax to be unconstitutional based on the Origination Clause. The Supreme Court ruled that the Senate amendment to insert the corpo-

¹⁸⁸ Story also very probably did not mean for courts to inquire into whether Congress really wanted to raise money by imposing taxes. In connection with protective tariffs, Story wrote:

If it be said, that the motive is not to collect revenue, what has that to do with the power? When an act is constitutional, as an exercise of a power, can it be unconstitutional from the motives with which it is passed? If it can, then the constitutionality of an act must depend, not upon the power, but upon the motives of the legislature No government on earth could rest for a moment on such a foundation. It would be a constitution of sand, heaped up and dissolved by the flux and reflux of every tide of opinion.

2 STORY, *supra* note 6, § 1086. This was the same point made by Judge Hough in *Hubbard v. Lowe*, under the Origination Clause. See *Hubbard v. Lowe*, 226 F. 135 (S.D.N.Y. 1915), *appeal dismissed mem.*, 242 U.S. 654 (1916).

¹⁸⁹ *Twin City Bank v. Nebeker*, 167 U.S. 196 (1897) at 203. The money was to pay for printing new currency that would be utilized by the banks themselves, and so that levy was arguably not traceable solely to the power to tax.

¹⁹⁰ 220 U. S. 107. In *Flint v. Stone Tracy* fifteen respondents challenged the Corporation Tax law of August 5, 1909. The case was hotly contested, during the era of the adoption and ratification of the Sixteenth Amendment. Oral arguments were heard on March 17th and 18th 1910 and again on January 17, 18 and 19, 1911.

rate tax for the inheritance tax was germane to the comprehensive House tax bill and therefore lawful:

“The bill having properly originated in the House, we perceive no reason in the constitutional provision relied upon why it may not be amended in the Senate in the manner which it was in this case. The amendment was germane to the subject-matter of the bill, and not beyond the power of the Senate to propose.”¹⁹¹

The corporate tax at issue before the Court in the bill had in fact been first considered in the House of Representatives in committee as part of the larger tax bill. However, the House committee drafting the bill

“had rejected the proposal because it did not think the revenue was needed and also because the Committee had decided already to propose an inheritance tax, as suggested by Taft at his inauguration. . . . the revenue bill that the House sent to the Senate for consideration consisted of the tariff provisions plus an inheritance tax but not a corporate tax.”¹⁹²

The Senate’s amendment of replacing the inheritance tax favored by the House with the corporate tax was a minor amendment to a comprehensive bill involving two alternatives already declared and considered in the House. The Court used this particular legislative action as an example of a Senate amendment to a revenue raising bill that was germane to the subject matter of the bill and therefore within the Senate’s amending power according to the Origination Clause. The Senate’s logic in reintroducing the corporate tax in the bill was to remove the criticism that the inheritance tax would be a “double-tax” (alongside the existing state inheritance taxes), and that another tax measure was necessary to replace the inheritance revenue measure originally favored by the House.¹⁹³ Thus the amendment was made in the Senate and the Court ruled it “germane to the subject matter of the bill” thereby establishing the standard for how far the Senate can amend a revenue raising bill.¹⁹⁴

In *Rainey v. United States*,¹⁹⁵ the Court again supported the power of the Senate to amend a House revenue raising bill. At issue this time was a Senate amendment imposing a tax on foreign-built pleasure yachts. The

¹⁹¹ 220 U. S. 107, 143.

¹⁹² Kornhauser, *supra* note 121, at 96. (citing 44 CONG. REC. 4717 (July 31, 1909)).

¹⁹³ *Id.* at 98.

¹⁹⁴ Interestingly to the legislative history of the bill and the modern House practice of blue-slipping, the House attempted to blue-slip a 1910 Senate amendment to amend the corporate tax at issue on origination grounds: “In 1910 the Senate added an amendment to the appropriation bill, H.R. 22643. . . . On April 1, 1910, the House discussed this amendment at length. Congressman Bartlett of Georgia immediately moved to return the bill to the Senate ‘with the request that the amendment be stricken from the bill, because it invades the constitutional privilege of the House to originate bills for the raising of revenue.’ After much discussion the resolution was rejected.” *Id.* at 127.

¹⁹⁵ 232 U.S. 310 (1914).

Senate's revenue amendment to the House's revenue bill was upheld, squarely adopting the lower court ruling. The Supreme Court did not address germaneness, declining to address the legislative purpose of the bill.

Both *Flint v. Stone Tracy* and *Rainey v. United States* leave unanswered the question of whether or not the Supreme Court has jurisdiction on Origination Clause grounds once an Act has been passed by both Houses of Congress. Questions of the scope and mechanics of the Origination Clause go unasked until the 1980s.¹⁹⁶ The Supreme Court took up the Origination Clause after 76 years of silence in *United States v. Munoz-Flores*.¹⁹⁷

In 1985 German Munoz-Flores was charged and convicted of a misdemeanor for aiding illegal immigrants in circumventing the immigration process. He was sentenced to probation and ordered to pay a "special assessment" of \$25 per count. The assessment was based on a federal statute, 18 U.S. C. 3013, in which courts imposed a monetary assessment for those convicted of a federal misdemeanor to be paid into the Crime Victims Fund, stipulated by the Crime Victims Act of 1984. Munoz-Flores appealed the assessment, arguing that the Crime Victims Act of 1984 violated the Origination Clause of the U.S. Constitution. Munoz-Flores reasoned that the Act was a revenue raising measure that originated in the Senate thereby violating the Origination Clause. The trial judge denied Munoz-Flores' motion and the District Court affirmed that decision. However the Court of Appeals for the Ninth Circuit reversed the decision on Origination Clause grounds while raising the political questions doctrine.

The Supreme Court heard oral argument on February 20, 1990. It focused on several questions: (1) was Origination Clause litigation a nonjusticiable political question; (2) if the Court did rule on the matter and ultimately strike an Act of Congress based on an Origination Clause violation, was this disrespectful to the legislative branch; and (3) is the statute a revenue raising bill? The Court answered all three of these questions in the negative. The Court ultimately ruled that the Crime Victims Act was not a revenue raising measure, and therefore did not violate the Origination Clause.

¹⁹⁶ Lower courts have followed the Supreme Court's decisions, finding very few violations of the Origination Clause. Lower courts have addressed whether or not the Origination Clause invokes the political questions doctrine, making the question outside of judicial review. There is a lively discussion of TEFRA, the Tax Equity and Fiscal Responsibility Act of 1982 in the lower courts. *See, e.g.* Texas Ass'n of Concerned Taxpayers, 772 F.2d 163 (5th Cir. 1985). The U. S. Supreme Court denied certiorari, docket 85-1262, cert denied 476 U.S. 1151, May 27, 1986. Also *see* Mulroy v. Block, 569 F. Supp. 256 (N.D.N.Y. 1983), *aff'd*, 736 F.2d 56 (2d Cir. 1984), cert. denied, 469 U.S. 1159 (1985). One of the most prominent lower court cases discussing the Origination Clause in general and germaneness in particular is *Hubbard v. Lowe* (S.D.N.Y. 1915).

¹⁹⁷ 495 U.S. 385, (1990). Again as noted above, lower courts discussed the interpretation of the Origination Clause even though the Supreme Court did not.

Marshall dismissed the political question claims following the logic of *Baker v. Carr*.¹⁹⁸ Courts can craft standards pertaining to bills for raising revenue and for where a bill originates: “Surely a judicial system capable of determining when punishment is ‘cruel and unusual’ when bail is ‘excessive’ ‘when searches are unreasonable,’ and when congressional action is ‘necessary and proper’ for executing an enumerated power is capable of making the more prosaic judgments demanded by adjudication of Origination Clause challenges.”¹⁹⁹

Marshall addressed the government’s argument that if the Court invalidated a law on Origination Clause grounds it would be disrespectful to the Congress. He explained that the judiciary is duty bound to review the constitutionality of congressional enactments. A purview of the judiciary, long recognized and held, Marshall dismissed the government’s argument out of hand. Justice Stevens in concurrence made the claim that Congress can better determine whether or not a bill violates the Origination Clause. Stevens argued that a bill that originates unconstitutionally can still be law if passed by both Houses of Congress and signed by the President. He reasoned that the House can easily defend its origination power by not agreeing to Senate amendments or voting in favor of legislation. Scalia in a separate concurrence also agreed that judicial deference to the legislative branch is preferable since Congress as a coequal branch can make Origination Clause determinations. Both Scalia and Stevens agreed that there is no Origination Clause violation since the Crime Victims Act is not a revenue raiser, but they differed from Marshall and the majority that courts should determine Origination Clause violations.

The Court concluded in *Munoz-Flores* that, “[t]he special assessment statute is not a ‘Bil[l] for raising Revenue’ and, thus, its passage does not violate the Origination Clause. This case fell squarely within the holdings of *Twin City Bank v. Nebecker* and *Millard v. Roberts* that a statute that creates, and raises revenue to support, a particular governmental program, as opposed to a statute that raises revenue to support government generally, is not a ‘Bil[l] for raising Revenue’”²⁰⁰:

The provision was passed as part of, and to provide money for, the Crime Victims Fund. Although any excess was to go to the Treasury, there is no evidence that Congress contemplated the possibility of a substantial excess, nor did such an excess in fact materialize. Any revenue for the general Treasury that § 3013 creates is thus incidental to that provision’s primary purpose.²⁰¹

To date, the only Supreme Court decision to articulate the Judicial Branch’s role in Origination Clause challenges is *United States v. Munoz-Flores* in 1990. The Court clarified the modern role of courts in Origina-

¹⁹⁸ 369 U.S. 186 (1962).

¹⁹⁹ 495 U.S. 385, 396 (1990).

²⁰⁰ 495 U.S. at 386-87.

²⁰¹ *Id.* at 387.

tion Clause claims: “A law passed in violation of the Origination Clause would thus be no more immune from judicial scrutiny because it passed by both Houses and signed by the President than would a law passed in violation of the First Amendment.”²⁰² Here, *Munoz* departed quite dramatically from the old Court standard regarding Origination Clause challenges expressed in *Field v. Clark Boyd* (1892)²⁰³ that the judiciary is bound to respect Congress’s indications of a Bill’s origination source via its formally enrolled status.

This departure from precedent in *Munoz* thus presents an opportunity to re-evaluate the deeper historical record and the few substantive Origination Clause Court decisions that exist, all of which indicates the need for a more careful consideration of this often overlooked constitutional provision.

The Court may be presented with its first post-*Munoz* Origination Clause case in this coming year in the *Sissel v. HHS* challenge against the Affordable Care Act. The district court judge in that case rejected the plaintiff’s Origination Clause challenge, ruling that the tax under consideration in the Affordable Care Act did not make it a “Bill for raising Revenue.”²⁰⁴ Further, the judge argued that even if the ACA were considered a “Bill for Raising Revenue” as the Constitution understood that category of legislation, the bill formally originated in the House and was germanely amended by the Senate.²⁰⁵ However, all three of these claims seem at least disputable. Based on our preceding research, the case appears far less conclusive than the district court judge made it out to be. First, it appears that the 17 Senate introduced taxes of the ACA very well may place it within the category of legislation intended by the original meaning of the phrase “Bills for raising Revenue.” Second, it is not clear that the vestigial bill number, “H.R.3590”, (when separated from any of the original bill’s substance) is sufficient to satisfy that the bill originated in the House of Repre-

²⁰² *Id.* at 397.

²⁰³ 143 U.S. 649.

²⁰⁴ *Sissel v. U. S. Dep’t of Health & Hum. Services*, No. 10-1263, slip op. at 14 (D.D.C. June 28, 2013) : “Congress’s preference would be for the individual mandate to raise zero revenues, and thus the provision cannot fairly be characterized as a ‘Bill[] for raising Revenue.’”

²⁰⁵ *Id.* at 22: “[E]ven assuming the individual mandate was a ‘Bill[] for raising Revenue,’ that bill ‘originated in the House of Representatives’ as H.R. 3590 and was later duly amended by the Senate in a manner consistent with the Origination Clause.” However, the plaintiff in the case maintains that the Court’s decision in *NFIB* alters the constitutional analysis of the ACA’s taxing provisions: “If the charge for not buying insurance is seen as a federal tax, then a new question must be asked,” said Pacific Legal Foundation Principal Attorney Paul J. Beard II. “When lawmakers passed the ACA, with all of its taxes, did they follow the Constitution’s procedures for revenue increases? The Supreme Court wasn’t asked and didn’t address this question in the *NFIB* case.” <http://www.pacificlegal.org/cases/Tax-raising-Affordable-Care-Act-started-in-wrong-house-of-Congress>

sentatives according to the Constitution's requirement. This does not appear justified under our historical review, and it is further questionable under Court precedent post-*Munoz*. Third, the history of the Clause does not seem to authorize (and the Supreme Court has never condoned) the use of the "gut and amend" parliamentary maneuver as within the Senate's power to amend revenue raising bills. (See preceding section of the "Original Public Understanding" of germaneness and the phrase "as on other Bills" in the context of *Sissel v. HHS*) For these among many other factors unique to the particular legislation and its legislative context beyond the scope of this paper, the case may present a unique Origination Clause challenge in the history of limited opportunities the Court has had to address the Clause. What appears to be in the government's favor in this case is the somewhat deferential standard the Court often took to Senate revenue measures in the 20th century prior to *Munoz*. What is not in the government's favor is the historical meaning of the Clause as we read it, and the distinctness of the current case from those in past Origination Clause rulings. The meaning of the Clause and the specific context of those cases as described above should be carefully considered before either extending to or distinguishing past Court precedent from a legislative project as enormous and significant as the ACA.

VII. THE COURT'S STANDARD, IMPLICATIONS, AND THEORETICAL CHALLENGES

While the Court has provided a narrow standard for what bills are considered revenue raising bills within the context of Article 1, §7, and if classed as a revenue raising bill, a general standard that any Senate amendments must be germane to the subject matter of the House originated bill, it has never addressed the questions of purpose and scale.

According to the Court's precedent through the cases presented to date, bills to which revenue raising is merely an incidental effect in the pursuit of some other legislative ends are not considered revenue raising bills. This standard was born out of a series of cases in which the taxes/fees imposed by Congress were of relatively small size and of narrow application as they attempted to exercise an enumerated constitutional power. In *Norton* the Congress was attempting to provide for a post system. In *Twin City v Nebeker* the Congress was attempting to provide for a currency. In *Millard*, the Congress was attempting to develop the District of Columbia. In *Flint*, the House was exercising its comprehensive taxing power and the Senate offered germane amendments to one element of that tax bill. In *Munoz*, Congress was attempting to enforce immigration law through criminal penalties. What if the Senate originated a taxing bill containing revenue raising provisions that were not a "means for effectively accomplishing" some enumerated "great object" as existed in all of the previous cases? Could the Senate originate a tax simply to exercise the taxing power and thereby claim that the revenue generation was merely incidental to the Congresses legitimate exercise of the taxing power? This seems to be the

unique and judicially novel question posed by the *Sissel* case given the *NFIB* ruling.

What is particularly problematic to this judicial standard is that it provides the Senate a blank check to originate any and all taxes it can couch as necessary to execute some other enumerated power. In theory, under this standard the entire federal budget and all receipts of the IRS could be designed and controlled through Senate originated bills. So long as the bills are compartmentalized and written to execute purposes other than revenue raising. Every tax could be labeled a “revenue offset” to the appropriation’s purpose contained in the Senate bill. This would circumvent and nullify any substantive meaning of Article 1, §7 of the Constitution.

While the Court has historically given the Senate considerable leeway to originate bills that do in fact tax people in small amounts for limited ends, it has never clearly addressed the questions “for what reasons,” and “how much” the Senate can tax. If the answer to those two questions is “any reason” and “any amount,” then there is no Origination Clause. The Court’s judicial interpretation of the clause throughout the 20th century remained relatively uncontroversial as the only cases it has upheld under Origination Clause scrutiny involved either germane amendments to large scale House originated tax bills, or else, relatively minor taxes that were truly incidental to a limited and clearly enumerated Senate legislative function.

VIII: CONCLUSIONS

The principle behind the Origination Clause was well established on the American continent during the 17th and 18th centuries. The perceived circumvention of the principle through various taxing measures instituted by Parliament in the 1760s was the primary ideological and legal argument for the Revolution. Of the nine States with bicameral legislatures by 1790, seven had lower house Origination Clauses. Most of them allowed upper house amendments of revenue raising bills. However, the belief that the mechanism of Senate amending power was meant by the Framers and the public they represented to allow the Senate “some” taxing power, misses the historical context and concerns of the clause. More often than not, the intent of adding a power for the Senate to amend money bills was to prevent an arrogating lower-house from tacking non-revenue raising measures onto revenue raising bills with the nefarious intent of circumventing the Senate’s legitimate input on the attached non-revenue raising matters. Ironically, the Court’s modern interpretation of the Origination Clause has favored Senatorial taxing power and enabled the opposite and equivalent abuse to be carried out by the Senate: The Senate now tacks “money bills” to unrelated House bills to circumvent the House’s constitutionally delineated prerogative of originating new taxes. The opposite primary fear intended to be avoided through an origination requirement was evidenced in the British experience with exclusive origination in the House of Commons as well as in States such as Maryland and Delaware. All added clear

amendment clauses as well as limitations on what constituted a “revenue raising bill” in order to prevent an abusive lower house. However, none of these were an attempt to “share” original taxing power with the upper house. To “originate” seems to still have meant the same understanding that was specified in the Maryland law of 1650 that all taxing measures must be “first had and declared in General Assembly.” By avoiding judicial activism on Congressional legislation, the Court seems to have passively permitted constitutional deviation from the Senate’s intended role throughout its 20th century precedent, thereby upsetting the intended role of the House of Representatives in controlling the purse strings of the people. If the Court were to uphold a new challenge containing a broader taxing measure on a larger scale than those previously upheld by the Court, it would highlight the difficulties of the Court’s standard and its divergence from the meaning of the Origination Clause.

In the Constitutional Convention, the Origination Clause was the primary bargaining chip used to bridge the disagreement between the large and small States that threatened progress on the Constitution. Politically, it was a concession made by the smaller States to limit the power of the Senate in exchange for allowing them to retain equal representation in that branch under the new Constitution. Philosophically and ideologically, it was a paramount expression and safeguard of democratic liberalism and popular sovereignty. While there were select concerns about it in the debates, both parties seemed to agree that the public would not support ratification of the Constitution if the popular clause were omitted. The public understanding of the clause was conveyed during the period of ratification by James Madison:

The House of Representatives cannot only refuse, but they alone can propose the supplies requisite for the support of the government. They in a word hold the purse; that powerful instrument by which we behold in the history of the British constitution, an infant and humble representation of the people, gradually enlarging the sphere of its activity and importance, and finally reducing, as far as it seems to have wished, all the overgrown prerogatives of the other branches of the government. This power over the purse, may in fact, be regarded as the most compleat and effectual weapon with which any constitution can arm the immediate representatives of the people, for obtaining a redress of every grievance, and for carrying into effect every just and salutary measure.²⁰⁶

²⁰⁶ THE FEDERALIST NO. 58 (James MADISON). That Madison is specifically referencing taxation here and not merely appropriations is evident from FEDERALIST NO. 45: “the present Congress have as complete authority to REQUIRE of the States indefinite supplies of money for the common defense and general welfare, as the future Congress will have to require them of individual citizens” Likewise, this plain understanding is evident from early House records: *see* Journal of the House, March 8th 1792: “Resolved, That the Secretary of the Treasury be directed to report to this House his opinion of the best mode for raising the additional supplies requisite for the ensuing year”. *See also* 3 MAX FARRAND, THE RECORDS OF THE FEDERAL CONVENTION OF 1787 at 356 (1911): James Madison speaking in the House of Representatives on 15 May, 1789:

Not much has changed since the ratification of the Constitution in terms of our theory of mixed legislatures. However, several key developments might give the Court pause:

1. *Direct election of Senators:* The 17th amendment could be argued to mitigate concerns over a lack of strict enforcement of the Origination Clause. The original indirect election of Senators was cited as one among several reasons why the House was more appropriate than the Senate for proposing taxing measures. However, all other “aristocratic” characteristics of the Senate (term lengths, non-proportional representation, non-local representation, etc.) remain the same today.

2. *Apportionment of taxes by population:* The 16th amendment could be argued to aggravate concerns and favor a strict enforcement of the Origination Clause. Joseph Story argued in the context of the appropriateness of the Senate’s amending power that “above all, as direct taxes are, and must be, apportioned among the states according to their federal population . . . there seems a peculiar fitness in giving to the senate a power to alter and amend, as well as concur with, or reject all money bills.”²⁰⁷ Whether labeled “direct” or “indirect”, the federal government no longer attempts to apportion any taxation according to the census.

3. *Ratio of Representatives to constituents:* Today’s America is of a far greater scale than that of 1787. With over 300 million Americans today, each congressman represents about 700,000 constituents. The framers sought to limit this ratio in the second half of the proposed Origination Clause “concession” to one representative for no more than 40,000 constituents. Although this requirement did not make it into the final draft of the Constitution, it was the first of the 12 proposed amendments for the bill of rights. It is the only un-ratified one today. The House of Representatives, let alone the Senate is far less representative of the “local knowledge” of the concerns and circumstances thought appropriate to exercise the power of proposing new taxes. This would aggravate concerns in favor of a strict enforcement of the Origination Clause.

Finally, and although not drawn out in this analysis, there are fundamental issues with Senate originated tax measures that strike at our Constitution’s basic theory of representation and the taxing power. As Madison

“The constitution, as had already been observed, places the power in the House of originating money bills. The principal reason why the constitution had made this distinction was, because they were chosen by the People, and supposed to be best acquainted with their interests, and ability. In order to make them more particularly acquainted with these objects, the democratic branch of the Legislature consisted of a greater number, and were chosen for a shorter period, so that they might revert more frequently to the mass of the People.”

²⁰⁷ 2 STORY, *supra* note 6, § 873, at 341.

noted in the Convention while reflecting on the consequences of granting the smaller States equal representation in the Senate, “[t]he majority of the States might still injure the majority of the people. . . . They could extort measures repugnant to the wishes & interests of the Majority [and] . . . could *impose* measures adverse thereto.”²⁰⁸ For this along with the many other reasons already mentioned, the Senate was never intended to write taxes and was explicitly forbidden from doing so in the Constitution.

Our analysis has reviewed the historical meaning and intent of the clause, Supreme Court precedent, and bicameral and federal theory considerations for future Court adjudication of Origination Clause challenges. From all three perspectives, there is reason for the Court to strengthen their enforcement of the Origination Clause. Specifically, if the Court does not enforce a germaneness requirement to Senate tax additions through amendment, then the Origination Clause will become wholly superfluous. If there is no germaneness standard to Senate tax additions then there is no substantive distinction between general legislative powers and the power to tax. The founders and the ratifying public understood such a distinction and expected it through article 1, §7. If this “distinction between legislation and taxation” is no longer protected and viewed as “essentially necessary to liberty”²⁰⁹ in the American tradition, then the nature of our Constitutional system of law has changed significantly since the voicing and codification of this establishing principle. The Senate was never entrusted with and was constitutionally forbidden from the original exercise of the taxing power.

Throughout the 20th century the Court developed a narrow standard for what bills are considered “revenue raising bills” within the context of Article 1, §7, and if classed as a revenue raising bill, a general standard that any Senate amendments must be germane to the subject matter of the House originated bill. While the somewhat passive evolution of this standard over the 20th century has survived as relatively uncontroversial given the small scale and the nature of the cases that gave rise to it, the Court will have to revisit the standard if broader challenges are presented in order to preserve any substantive meaning and effect of the clause in the Constitution, the Revolutionary principle, and our theory of mixed legislatures. The recent string of cases challenging the Senate origins of the vast majori-

²⁰⁸ MADISON, *supra* note 11, at 224.

²⁰⁹ William Pitt, *On an address to the Throne, in which the right of taxing America is discussed*, 17 Dec., 1765 reprinted in THE TREASURY OF BRITISH ELOQUENCE (compiled by Robert Cochrane, London and Edinburgh, 1877) at 140-141. *See also* United States v Munoz-Flores, 495 U.S. 385,395, 397 (1990):

Provisions for the separation of powers within the Legislative

Branch are thus not different in kind from provisions concerning relations between the branches; both sets of provisions safeguard liberty. . . . A law passed in violation of the Origination Clause would thus be no more immune from judicial scrutiny because it was passed by both Houses and signed by the President than would be a law passed in violation of the First Amendment.

ty of the Affordable Care Act's taxes, including *Sissel v. HHS*, may force the Court into the uncomfortable position of deciding between the survival of the signature healthcare reform law and the survival of some minimal degree of substance in and force behind the Constitution's Origination Clause.

Appendix

Statements from the Ratification Debates Evincing the Public Understanding of the Clause:

- “They [the Senate] may restrain the profusion of errors of the house of representatives, but they cannot take the necessary measures to raise a national revenue.”²¹⁰
- “Without their [House of Representatives] consent no monies can be obtained, no armies raised, no navies provided. They alone can originate bills for drawing forth the revenues of the Union, and they will have a negative upon every legislative act of the other house. So far, in short, as the sphere of federal jurisdiction extends, they will be controllable only by the people, and in contentions with the other branch, so far as they shall be right, they must ever finally prevail. Such, my countrymen, are some of the cautionary provisions of the frame of government your faithful Convention have submitted to your considerations.”²¹¹
- “. . . and in the House of Representatives must all money bills originate.”²¹²
- “Let U.S. examine how far the peculiar constitution of our federal Senate will give U.S. the advantages of the second legislative branch without subjecting U.S. to the dangers usually apprehended from such bodies.”²¹³
- “Answer [to objections of tyranny in laying and collecting taxes]: Who are the members that constitute this [House of Reps.] body – the *people* or their representatives? Can they do any act that they themselves are not bound by; and if they lay *excessive taxes*, the people will

²¹⁰ An American Citizen II: On the Federal Government *Philadelphia Independent Gazetteer*, 28 Sept. 1787 reprinted in DHRC, *supra* note 97, V.13 at 265: Reprinted 21 Oct. 1787 under title: “On the safety of the people, from the restraints imposed on the Senate.” V.2. at 143.

²¹¹ Trench Coxe, An American Citizen III, *Independent Gazetteer*, 29 Sept. 1787 reprinted in DHRC, *supra* note 97, V.2. at 145.

²¹² One of the People, PENNSYLVANIA GAZETTE, 17 Oct. 1787 reprinted in DHRC, *supra* note 97, V.2 at 191.

²¹³ A Democratic Federalist, *Independent Gazetteer*, 26 Nov. 1787 reprinted in DHRC, *supra* note 97, V.2 at 294-95.

have it in their power to return other men (vide section 7th of 1st [Article] for the origination of *revenue bill*).”²¹⁴

- “The two branches will serve as checks upon the other; they have the same legislative authorities, except in one instance. Money bills must originate in the House of Representatives.”²¹⁵
- “Some of the powers of the Senators are not with me the favorite parts of it, but as they stand connected with the other parts, there is security against the efforts of that body. It was with great difficulty that security was obtained, and I may risk the conjecture, that if it is not now accepted, it never will be obtained again from the same states. Though the Senate was not a favorite of mine, as to some of its powers, yet it was a favorite with the majority of the Union, and we must submit to the majority, or we must break up the Union.”²¹⁶ (Referencing the deal of the Great Compromise)
- “The Senate are incapable of receiving money, except what is paid to them out of the public treasury. They cannot vote to themselves a single penny, unless the proposition originates from the other house.”²¹⁷
- “Sir, there is another principle that I beg leave to mention. Representation and direct taxation, under this Constitution, are to be according to numbers.”²¹⁸
- “Wars are inevitable, but war cannot be declared without the consent of the immediate Representatives of the people; there must also originate the law which appropriates the money for the support of the army, yet they can make no appropriation for longer than two years.”²¹⁹
- “Passing over my lesser matters, I proceed to section 7, which is in these words ‘All bills for raising revenue. . . but the SENATE may propose or concur with amendments as on other bills.’ I would here only observe that the Commons of Great Britain will not suffer the House of Lords to make the least alteration in a money bill; however,

²¹⁴ The Pennsylvania Convention Debates, 28 Nov. 1787 *reprinted in* DHRC, *supra* note 97, V.2 at 411.

²¹⁵ The Pennsylvania Convention Debates, 1 Dec. 1787 *reprinted in* DHRC, *supra* note 97, V.2 at 451.

²¹⁶ Proceedings and Debates of the [Pennsylvania] Convention, 4 Dec. 1787 *reprinted in* DHRC, *supra* note 97, V. 2 at 480.

²¹⁷ *Id.* at 490.

²¹⁸ (James Wilson) Proceedings and Debates of the [Pennsylvania] Convention, 4 Dec. 1787 *reprinted in* DHRC, *supra* note 97, V.2 at 497.

²¹⁹ Proceedings and Debates of the [Pennsylvania] Convention, 10 Dec. 1787 *reprinted in* DHRC, *supra* note 97, V.2 at 528-39.

the Crown has found means to corrupt a sufficient number of the Commons to draw forth the blood and treasure of the nation.”²²⁰

- [in response]“that this proves that the framers of the Constitution were no servile imitators of the British theory of government, nor under the special influence of Mr. [John] Adams’s sentiments, for the ‘British House of Commons will not suffer the House of Lords to make the least alteration in a money bill.’ . . . The body of the people must be convinced that the purse of the nation will be as safe in the hands of their Representatives in Congress as of their representatives in the state assemblies.”²²¹
- “The admission however of the smaller States to an equal representation in the Senate, never would have been agreed to by the Committee or by myself as a member of it without the provision ‘that all bills for raising or appropriating money & for fixing the salaries of the officers of Government’ should originate in the house of Representatives & ‘not be altered or amended’ by the Senate ‘& that no money should be drawn from the treasury’ ‘but in pursuance of such appropriations’.”²²²
- “It is objected, that it is dangerous to allow the senate a right of proposing alterations or amendments in money-bills – that the senate many by this power increase the supplies and establish profuse salaries – that for these reasons the lords in the British parliament have not this power, which is a great security to the liberties of Englishmen. I was much surprised at hearing this objection, and the grounds upon which it was supported. . . . But every supposed control the senate by this power may have over money-bills, they can have without it, for by private communications with the representatives, they may as well insist upon an increase of the supplies, or salaries, as by official communication.”²²³
- “The 7th section of the first article in the proposed constitution says, ‘All bills for raising revenue shall originate in the House of Representatives; but the Senate may propose or concur with amendments as on other bills.’ This is putting a power into the hands of the Senate with more safety than can generally be done – it is giving them the power of doing good, almost without the possibility of doing harm; for it would be folly to suppose that the House of Representatives, or any other body of men, could form a bill so completely perfect in all its

²²⁰ Letter from Massachusetts and Letter from New York, *Connecticut Journal*, 17 Oct. 1787 reprinted in DHRC, *supra* note 97, V.3 at 377.

²²¹ Letter from New York, 24 Oct. 1787 reprinted in DHRC, *supra* note 97, V.3 at 385.

²²² Elbridge Gerry, Defense of Elbridge Gerry’s Actions in the Massachusetts Convention on 19 Jan. 1788 reprinted in DHRC, *supra* note 97, V.6 at 1269-70.

²²³ Massachusetts Convention Debate, 23 Jan. 1788 reprinted in DHRC, *supra* note 97, V.6 at 1327.

parts as to admit of no amendment. A revenue bill will now have a double chance of attaining to perfection. The House of Representatives will discuss, form and send it up – the Senate will have it in their power to deliberate, debate upon it, and propose amendments, if necessary; but they can go no further.”²²⁴

- “The senate has the power of originating all bills, except *revenue bills*, in common with the house of representatives. . . From this exclusion of the *senate* with respect to *money bills*, it is plain, that this body does not possess such extensive legislative power, as the house of representatives.”²²⁵
- “3d. ‘The senate have the power of altering money bills;’ and why not? Because the Lords in England, an hereditary aristocracy, have not, of late years, been permitted by the commons to exercise this power, shall the senate, a rotatory body, chosen by the representatives of the people, be deprived of this essential right of legislation? The people cannot be taxed, but, by the consent of their immediate representatives.”²²⁶
- “In this the Constitution is an improvement upon that of England: There all money bills must not only originate but must be perfected in the House of Commons: Here though the Senate cannot originate such bills, yet they have the power of amending them, and by that means have an opportunity of communicating their ideas to the House of Representatives upon the important subject of taxation.”²²⁷
- “So able an advocate as you for the new constitution may also assign a good reason why Delaware, that pays but a sixty seventh part of the general expenses, should vote on a money bill in the senate equally with Virginia that pays a sixth part of the same expenses? Perhaps you may satisfy U.S. that another convention cannot be obtained to remedy the defects that are so apparent in the proposed system.”²²⁸
- “Two years are the utmost time for which the money can be given. It will be under all the restrictions which wisdom and jealousy can sug-

²²⁴ Brutus *Virginia Journal*, 6 Dec. 1787 reprinted in DHRC, *supra* note 97, V.8 at 214-15.

²²⁵ Valerius [to Richard Henry Lee], *Virginia Independent Chronicle*, 23 Jan. 1788 reprinted in DHRC, *supra* note 97, V.8. at 316.

²²⁶ Civic Rusticus, *Virginia Independent Chronicle*, 30 Jan. 1788 reprinted in DHRC, *supra* note 97 V.8 at 335.

²²⁷ A Native of Virginia: Observations upon the Proposed Plan of Federal Government, Apr. 2 1788 reprinted in DHRC, *supra* note 97, V.9 at 668.

²²⁸ Brutus, *Virginia Independent Chronicle*, 14 May 1788 reprinted in DHRC, *supra* note 97, V.9 at 802.

The Origination Clause

gest, and the original grant of the supplies must be made by the House of Representatives, the immediate representatives of the people.”²²⁹

²²⁹ An American Citizen IV: On the Federal Government Philadelphia, 21 Oct. 1787 reprinted in DHRC, *supra* note 97, V.13 at 435.

FROM *LARK RISE* TO *THE STORIED CITY*

John Martinez*

S.J. Quinney College of Law at the University of Utah

ABSTRACT

This article suggests that storytelling can be a useful tool for analyzing local government authority. Using a cognitive science approach to law, the article critically evaluates conventional legal analysis of local government authority in comparison to a storytelling approach. The article compares conventional legal analysis and storytelling in three illustrative settings: local government condemnation of private property for resale to a private developer, delegation of land use authority to neighborhood groups, and local government attempts to zone out nontraditional families. That comparison reveals that storytelling can help illuminate what is at stake in the interpretation of local government authority, and perhaps help us to arrive at better solutions.

CONTENTS

I.	INTRODUCTION	142
II.	A COGNITIVE SCIENCE APPROACH TO LAW.....	143
III.	LOCAL GOVERNMENT AUTHORITY.....	145
IV.	UNPACKING CONVENTIONAL LEGAL ANALYSIS OF LOCAL GOVERNMENT AUTHORITY.....	148
V.	THE UNSTATED ASSUMPTIONS OF CONVENTIONAL LEGAL ANALYSIS OF LOCAL GOVERNMENT AUTHORITY.....	153
VI.	A COGNITIVE CRITIQUE OF CONVENTIONAL LEGAL ANALYSIS	158
	A. MICHELMAN’S APPROACH AS METAPHOR: “PYRAMID” THINKING.....	158
	B. THE ORDER WE SEE IS THE ORDER WE IMPOSE	159

VII. STORIES AS ALTERNATIVE STRATEGIES FOR DEFINING LOCAL GOVERNMENT AUTHORITY.....	160
A. THE TELLING POINT OF STORIES	160
B. EXAMPLES OF ISSUES OF LOCAL GOVERNMENT POWER CONSIDERED AS STORIES	161
VIII. CONCLUSION	164

I. INTRODUCTION

In *Lark Rise to Candleford*, Flora Thompson, through her character Laura Timmins, tells stories about the daily struggles of the residents of a poor English hamlet coping with the wrenching changes of late 19th Century England.¹ Cooperative working of open fields which had supported self-sufficiency is replaced by enclosed landholdings, industrialization and imminent starvation. The contrasts between the stories of Queenie, the beekeeper and Dorcas Lane, the postmistress, reveal the evolution between an agrarian communal society and an industrialized economy. Their stories tell so much more than conventional legal description could provide.

This article proposes *The Storied City* as a better way to understand what is at stake in regard to local government authority and to help us to arrive at better solutions. The article describes a cognitive science approach to law, uses it to critically evaluate conventional “pyramid” legal analysis of local government authority, and suggests stories as alternative models for defining such authority. The article illustrates the storytelling analytical approach in three situations: a local government's condemnation of private property for resale to a private developer, the delegation of land use control authority to neighborhood groups, and local government attempts to zone out nontraditional families.

Part I describes a cognitive science approach to law. Part II traces the evolution of local government powers and focuses on the particularly intractable problem of adequately defining their proper scope. Part III unpacks conventional local government legal analysis by exploring the writings of Frank Michelman, a prominent local government law theorist. Part IV reveals the unstated assumptions of conventional legal analysis and the

* Professor of Law, S.J. Quinney College of Law at the University of Utah. This article was funded in part by the University of Utah College of Law Excellence in Teaching and Research Fund. I would like to thank my wife Karen Martinez for all her encouragement and support.

¹ FLORA THOMPSON, *LARK RISE TO CANDLEFORD: A TRILOGY* (1945). A television series from Flora Thompson's trilogy of novels produced by the BBC aired from 2008 to 2011. For the listing of the program episodes, see <http://www.bbc.co.uk/programmes/b00gbb10/episodes/guide> (last visited Sept. 30, 2013).

serious and unsettling consequences of those assumptions for defining local government authority. Part V reconstructs conventional local government legal analysis in cognitive terms and Part VI concludes by suggesting *stories* as alternative cognitive strategies for defining local government authority.

II. A COGNITIVE SCIENCE APPROACH TO LAW

Cognitive science² is the study³ of how we acquire, process and use information.⁴ A cognitive model may be thought of as a lens through which we see the world, or as a template we use to make sense of what we perceive. The insight of cognitive science is that meaning is in large part a function of our use of cognitive models to organize the world around us.⁵ These cognitive tools embody our experiences, and are acquired, stored and modified through use.⁶ A significant body of scholarship has developed around the application of cognitive science to law.⁷

² This description of cognitive science is drawn in part from my previous article in the area: John Martinez, *A Cognitive Science Approach to Teaching Property Rights in Body Parts*, 42 J. LEGAL EDUC. 290 (1992). I have also suggested a cognitive science approach to takings jurisprudence in my Government Takings treatise. See JOHN MARTINEZ, GOVERNMENT TAKINGS (Thomson-Reuters/West, "GOTA" database on WestLaw) Ch. 16 (Cognitive Science Approach to Takings: Thinking Outside the Box).

³ Cognitive science is not a single unified field, but instead exists as a composite of portions of philosophy, psychology, artificial intelligence, linguistics, anthropology, and neuroscience. HOWARD GARDNER, THE MIND'S NEW SCIENCE: A HISTORY OF THE COGNITIVE REVOLUTION 6-7, 38-45 (1985) (hereinafter GARDNER).

⁴ *Id.* at 6. There is a substantial body of scholarship on cognitive science. See, e.g., PHILLIP J. DAVIS & REUBEN HERSH, THE MATHEMATICAL EXPERIENCE (1980); HOWARD GARDNER, THE MIND'S NEW SCIENCE: A HISTORY OF THE COGNITIVE REVOLUTION (2d ed. 1987); PHILIP NICHOLAS JOHNSON-LAIRD, THE COMPUTER AND THE MIND, AN INTRODUCTION TO COGNITIVE SCIENCE (1988); ZENON W. PYLYSHYN, COMPUTATION AND COGNITION, TOWARD A FOUNDATION FOR COGNITIVE SCIENCE (1984); STEPHEN P. STICH, FROM FOLK PSYCHOLOGY TO COGNITIVE SCIENCE: THE CASE AGAINST BELIEF (1983). See generally Raffaele Caterina, *Comparative Law and the Cognitive Revolution*, 78 TUL. L. REV. 1501 (2004).

⁵ There is indeed structure in the environment, but except at fairly simple perceptual levels, it must be learned through experience. When it is learned it becomes a mental structure that guides the course of future information extraction. The knowledge that is so gained does not consist of lists of unrelated factors or a heap of haphazard associations. As Piaget so often emphasized, the mind has a tendency to organize itself. JEAN MATTER MANDLER, STORIES, SCRIPTS AND SCENES: ASPECTS OF SCHEMA THEORY 113 (1984).

⁶ GARDNER, *supra* note 3, at 30; Edward L. Rubin, *The Practice and Discourse of Legal Scholarship*, 86 MICH. L. REV. 1835, 1838-39 (1988) ("(O)ur sense of reality is determined ... by the *cognitive constructs* that make thinking possible.") (emphasis added). ... Any number of vocabularies and conceptual frameworks have been constructed in an effort to characterize the representational level -- scripts, schemas, symbols, frames, images, mental models, to name just a few.

Cognitive science posits that the process of formation, storage, use and transformation of cognitive models is systematic and imaginative.⁸ The first step involves the “basic experiences” common to all human beings by virtue of the general structure and functioning of the human organism in its environment.⁹ For example, one of these basic experiences is our primal discovery that we can obtain desired objects by moving toward them through space.¹⁰ As a second step, we conceptualize—or imagine—cognitive models that embody these basic experiences. Thus, we may imagine a source-path-goal cognitive model to represent the basic experience of obtaining a desired object by moving toward it through space.¹¹ Third, we may use the source-path-goal cognitive model and its source basic experience to “experience” more abstract purposive behaviors, such as our conceiving of a half-completed task as “being halfway there.”¹² Finally, we may use the source-path-goal cognitive model as the basis for metaphors to

GARDNER, *supra* note 3, at 383-84. See also MANDLER, *supra* note 6 (discussing the relationship of “schema theory,” another branch of the cognitive science tree).

⁷ See, e.g., Jay M. Feinman, *The Jurisprudence of Classification*, 41 STAN. L. REV. 661 (1989); Gerald P. Lopez, *Lay Lawyering*, 32 UCLA L. REV. 1 (1984); Pierre J. Schlag, *Missing Pieces: A Cognitive Approach to Law*, 67 TEX. L. REV. 1195 (1989); see also Timothy P. Terrell, *Flatlaw: An Essay on the Dimensions of Legal Reasoning and the Development of Fundamental Normative Principles*, 72 CALIF. L. REV. 288 (1984) (a delightful application of cognitive principles to explain legal reasoning by means of graphic representations, such as lines, planes and cubes). For an excellent discussion how entrenched cognitive models act as cognitive blocking mechanisms on counterhegemonic methods of legal analysis, see MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES 270-356 (1987) (toward cognitive theory of legitimation).

⁸ I am indebted to Steven L. Winter and Pierre J. Schlag for their pioneering work in the development of a cognitive approach to law. See Pierre J. Schlag, *Missing Pieces: A Cognitive Approach to Law*, 67 TEX. L. REV. 1195 (1989); Steven L. Winter, *Bull Durham and the Uses of Theory*, 42 STAN. L. REV. 639 (1990) (hereinafter Winter, *Bull Durham*); Steven L. Winter, *The Cognitive Dimension of the Agon Between Legal Power and Narrative Meaning*, 87 MICH. L. REV. 2225 (1989) (hereinafter Winter, *Agon*); Steven L. Winter, *Transcendental Nonsense, Metaphoric Reasoning, and the Cognitive Stakes for Law*, 137 U. PA. L. REV. 1105 (1989) (hereinafter Winter, *Transcendental Nonsense*); Steven L. Winter, *The Metaphor of Standing and the Problem of Self-Governance*, 40 STAN. L. REV. 1371 (1988) (hereinafter Winter, *Standing*). See also Martha Minow, *The Supreme Court, 1986 Term--Foreword: Justice Engendered*, 100 HARV. L. REV. 10 (1987) (exploring how Supreme Court Justices' unstated cognitive structures affect decisions). See generally Douglas W. Maynard, *Narratives and Narrative Structure in Plea Bargaining*, 22 LAW & SOC'Y REV. 449-81 (1988) (narrative as an alternative rhetoric for dispute resolution); Albert J. Moore, *Trial by Schema: Cognitive Filters in the Courtroom*, 37 UCLA L. REV. 273 (1989) (jurors' use of schemas to process evidence at trial).

⁹ Winter, *Transcendental Nonsense*, *supra* note 8, at 1133.

¹⁰ *Id.* at 1132.

¹¹ *Id.* Winter sets out some of the most important schemas as source-path-goal, container (in-or-out orientation), part-whole, front-back, center-periphery and balancing. *Id.* at 1147. These will be discussed later in the text as the experientialist epistemological approach is applied to legal doctrine.

¹² *Id.*

structure other aspects of our existence. Thus, the metaphor “Life is a Journey,” may be elaborated from the source-path-goal cognitive model, to conceptualize life.¹³

We use cognitive models and their derivative metaphors to understand, retain and apply more complex concepts¹⁴. Legal concepts can thus be profitably explored in terms of a cognitive science approach¹⁵. Conceived in terms of cognitive models, legal doctrine thus may be said to arrange legal analysis into core areas of certainty -- in which “most legally trained observers committed to applying [a] rule will experience the rule as having sufficient structure to constrain decision”¹⁶-- and peripheries, where the degree of “fit” between the cognitive model and the particular circumstances leads to indeterminacy, and where metaphoric extensions of the cognitive model inform the manner in which we resolve cases.¹⁷

III. LOCAL GOVERNMENT AUTHORITY

The nature and extent of powers held by local governments in America have been evolving over time.¹⁸ Moreover, the sources of governmental

¹³ *Id.* at 1115.

¹⁴ *Id.* at 1134.

¹⁵ In fact, “(e)xperientialist epistemology suggests that the courts will not function in a substantially different manner (if they self-consciously acknowledge the figurative discourse that clothes their decisionmaking): They will not be able to purge metaphors from their analyses, but will be driven to other metaphors.” *Id.* at 1164. “Metaphor is inevitable in legal analysis because it is central to human rationality; it is the primary mode of comprehension and reasoning.” *Id.* at 1166.

¹⁶ This same notion is expressed by Kent Greenawalt as a “determinate” character of legal rules:

The main criterion for judging the existence of a determinate answer is whether virtually any intelligent person familiar with the legal system would conclude, after careful study, that the law provides that answer.

Kent Greenawalt, *How Law Can Be Determinate*, 38 UCLA L. REV. 1, 3 (1990).

¹⁷ Winter, *Transcendental Nonsense*, *supra* note 8, at 1182-83. See also W. Quine, *Two Dogmas of Empiricism*, in WILLIAM R. BISHIN & CHRISTOPHER D. STONE, LAW, LANGUAGE, AND ETHICS, AN INTRODUCTION TO LAW AND LEGAL METHOD 326 (1972) (referring to rules as force fields with cores and peripheries); Duncan Kennedy, *Freedom and Constraint in Adjudication: A Critical Phenomenology*, 36 J. LEGAL EDUC. 518 (1986) (also referring to “force fields” of legal doctrines).

The rhetorical technique of conceptualizing legal analysis dealing with “force fields” with cores and peripheries, of course, is itself a cognitive model. It allows us to ask whether there are any rules for determining which particular legal rule-qua-cognitive model is applicable in any situation. This “characterization” step can have a powerful effect on the outcome in any particular case. The central project of this article, however, is not to deal with the substance of outcomes, but to demonstrate that the cognitive approach can help illuminate the way in which we do legal analysis.

¹⁸ An excellent book on governmental property power through the study of the city of New York is HENDRIK HARTOG, PUBLIC PROPERTY AND PRIVATE POWER, THE CORPORATION OF THE CITY OF NEW YORK IN AMERICAN LAW, 1730-1870 (1983) (hereinafter HARTOG). See also Gerald E. Frug, *Property and Power: Hartog on the Legal History of New York City* (Book Review), 1984 AM. BAR FOUNDATION RESEARCH J. 673; Carol M.

power and the manner in which we seek to control that power have developed in dynamic interaction with each other.¹⁹ Early conceptions of governmental authority rested almost exclusively on property.²⁰ Thus, kings formerly owned all property and merely shifted property entitlements around in order to effectuate royal policy. There was no “regulatory” sovereign power in the sense in which we usually understand governmental action today. This was also true of medieval towns, in which there was no separation between a town's property rights and its sovereignty rights.²¹

This property-based governance persisted in Eighteenth Century America, when local governments relied almost exclusively upon their property interests for social control. In his account of the evolution of the powers of New York City, Hendrik Hartog describes how that city's property and governmental rights were initially “blurred and mixed.”²² The original charter by Governor John Montgomerie in 1730 conferred governance and property powers on the city. The governance powers were the

Rose, *Public Property, Old and New* (Book Review), 79 NW. U. L. REV. 216 (1984); Joan C. Williams, *Review--The Development of the Public Private Distinction in American Law* (Book Review), 64 TEX L. REV. 225 (1985); Joan C. Williams, *The Invention of the Municipal Corporation: A Case Study in Legal Change* (Book Review), 34 AM U. L. REV. 369 (1985).

¹⁹ The most significant account of the importance of the role of forms of judicial review of local governmental action is by Professor Michelman, who suggests that two contradictory and ultimately irresolute, “judicial mentalities” exist in judicial elaboration of the public purpose and delegation doctrinal areas: an economic “public choice” model and a non-economic “public interest” or “community self-determination” model. Frank I. Michelman, *Political Markets and Community Self-Determination: Competing Judicial Models of Local Government Legitimacy*, 53 IND. L.J. 145, 147-48 (1977-78) (hereinafter Michelman, *Political Markets*).

Hendrik Hartog explores this dimension to some degree as well, explaining judicial refusal to recognize inherent right to local self-government as follows:

... Like their evangelical counterparts, judges distrusted the ability of public authorities (both legislators and city officials) to choose wisely. Their insistence on the autonomy of the law can be viewed as one way of distinguishing themselves from the political apparatus of the state, as, conversely, a way of identifying with the goals of moral reform. The doctrines they developed, although formally committed to legislative sovereignty, were shaped to ensure that courts would determine the legal consequences of public action.

From this vantage point, Dillon's Rule becomes an appropriate moral gesture, a way of compelling the legislature to take responsibility for the actions of an errant child. The city was not to be set loose on the streets of public action and expenditure freed from the constraints of its parent. The law would compel the legislature to superintend its charge. HARTOG, *supra* note 18, at 262-63.

²⁰ This was usually in the form of real estate, the quintessential form of productive wealth, but other forms of productive wealth, such as monopolistic rights, were used as well. See Rose, *supra* note 18, at 219 (“[F]irst, create an exclusive and lucrative proprietary right to perform some activity, then attempt to hedge the right with directions for its use, and finally grant it to someone and hope for the best.”).

²¹ Gerald E. Frug, *The City as a Legal Concept*, 93 HARV. L. REV. 1057, 1083-87 (1980).

²² HARTOG, *supra* note 18, at 21.

now-familiar police powers,²³ including the authority to pass regulations for the public good, run jails and operate courthouses. The property powers, however, included ownership of most of Manhattan Island north of Canal Street, the underwater land that surrounded Manhattan, and the Brooklyn waterfront. The city used its property power to raise revenue, for example, through conveying waterlots to wealthy citizens on condition that the purchasers build adjacent streets and docks.²⁴ Apparently, New York did not similarly use its property power to “regulate” in the modern sense, perhaps in part because regulation of conduct—directly through the exercise of sovereign power or indirectly through the use of property power—does not seem to have developed in political theory by that time.²⁵ Instead, cities concentrated on property management, which only indirectly or derivatively regulated conduct.²⁶

Primitive governmental power based on governmental property ownership atrophied with the reduction of government property.²⁷ Instead, governments came to rely more and more on their “sovereign” powers to tax, license and regulate private conduct directly as the means for effectuating social policy.²⁸

The emergence of generalized governmental power separate from governmental property power forced examination of the ways in which such power should be contained. John Dillon argued that control of city power through judicial supervision of that control was essential, but his proposed solution was a miserly interpretation of grants of state power over cities:

The courts, too, have duties, the most important of which is to require these corporations, in all cases, to show a plain and clear grant for the

²³ Police power represents the authority to regulate for the benefit of the general health, safety, welfare and morals. See *Village of Euclid v. Ambler Realty Co.*, 272 U.S. 365 (1926) (zoning).

²⁴ HARTOG, *supra* note 18, at 68.

²⁵ *Id.* at 43.

²⁶ *Id.* at 40 (“The proper business of the (city) corporation was the management, care, and disposal of the real estate it owned.”). Significantly, the city of Philadelphia passed only one new ordinance between 1740 and 1776. JON C. TEAFORD, *THE MUNICIPAL REVOLUTION IN AMERICA: ORIGINS OF MODERN URBAN GOVERNMENT* 56-59 (1975).

²⁷ Professor Hartog suggests it might have been the other way around with respect to New York City. He argues that there was a generalized change in social consensus that local governments should simply be creatures of the state, that judicial doctrines for reviewing local governmental power adopted this consensus through strictly restricting local government power to that delegated from the state, and that thereafter New York City officials transferred away much of the city's remaining property, with no further thought about revenue or social coercion. HARTOG, *supra* note 18, at 7-8. We need not resolve which was the cause and which the effect, but merely recognize that property ownership as a source of local governmental power independent of state control became unavailable.

²⁸ See generally 1 JOHN MARTINEZ, *LOCAL GOVERNMENT LAW* (2d ed. 2012) (Thomson-Reuters/West, “LOGOVTLAW” database on WestLaw) §13:1 *et seq.*, (hereinafter, *LOCAL GOVERNMENT LAW*) (discussing various sources and limitations of local government authority).

authority they assume to exercise; to lean against constructive powers, and, with firm hands, to hold them and their officers within chartered limits.²⁹

This perspective resulted in the development of the Dillon Rule, whereby local governments were not deemed to have the power to engage in any activity unless such power was granted expressly or by necessary implication from grants of power from the state legislature.³⁰ Early cases turned on interpreting fairly general grants of power to local governments from state legislatures, such as the general police power to engage in activities for the health, safety, welfare and morals of the community.³¹ Courts infused meaning into these general grants of power through equally vague notions of what a “public purpose” is or what is “ultra vires” of a general grant of power.³² However, courts have questioned the Dillon Rule and have instead been willing to construe general grants of power from the legislature more generously or, in the alternative, have completely abandoned the Rule in favor of a presumption that local governments have power to engage in their activities unless they are prohibited from doing so by state legislation.³³

IV. UNPACKING CONVENTIONAL LEGAL ANALYSIS OF LOCAL GOVERNMENT AUTHORITY

Frank Michelman has written extensively regarding the evolution of local governmental authority³⁴ and has made two important contributions

²⁹ John F. Dillon, TREATISE ON THE LAW OF MUNICIPAL CORPORATIONS ch. 1, §9, at 25-26 (1872) (emphasis in original).

³⁰ See, e.g., Taxpayers and Citizens of the City of Foley v. City of Foley, 527 So.2d 1261 (Ala. 198) (adhering to the Dillon Rule); Borgelt v. City of Minneapolis, 271 Minn. 249, 250, 135 N.W.2d 438, 440 (1965) (applying Dillon Rule). For the argument that the Dillon rule is actually a rule of contextuality and flexibility, see 3 LOCAL GOVERNMENT LAW, §13:5.

³¹ See, e.g., Union Ice & Coal Co. v. Town of Ruston, 135 La. 898, 66 So. 262 (1914) (police power held to be power to regulate, not to go into business operating an ice plant in competition with private ice company).

³² See generally Michelman, *Political Markets*, supra note 19 (discussing local government power under “public purpose” doctrine). For a discussion of other common law concepts used to define local government powers in the context of municipal contracts, see Janice C. Griffith, *Local Government Contracts: Escaping from the Governmental/Proprietary Maze*, 75 IOWA L. REV. 277 (1990).

³³ See, e.g., State v. Dirnberger, 152 Minn. 44, 187 N.W. 972 (1922). For a careful analysis of the reasons to abandon the Dillon Rule, see State v. Hutchinson, 624 P.2d 1116 (1981). Professor Frug, in his path breaking work on cities, contrasts city powerlessness with the power of private corporations and suggests that cities should be given more of the powers possessed by private corporations. Gerald E. Frug, *The City as a Legal Concept*, 93 HARV. L. REV. 1057, 1062-67 (1980).

³⁴ See, e.g., Frank I. Michelman, *Law's Republic*, 97 YALE L.J. 1493 (1988) (hereinafter Michelman, *Law's Republic*); Frank I. Michelman, *The Supreme Court 1985 Term, Foreword: Traces of Self-Government*, 100 HARV. L. REV. 1 (1986); Frank I. Michelman, *Poli-*

to the field. First, he has described the conventional legal analytical framework for considering these questions. Second, his framework describes the critical role that standards of judicial review play in such analysis.

According to Michelman,³⁵ analysis of legal issues implicates an organizational structure which at its most specific considers the factual circumstances giving rise to the issues and at its most general considers philosophical conceptions of value. The entire organizational framework asks the following questions:

- (1) Factual Circumstances: What factual circumstances give rise to the specific problems involved?
- (2) Doctrine(s): What established legal doctrines would ordinarily be expected to resolve the questions thereby posed?
- (3) Standards of Judicial Review: What standards of judicial review would a court be expected to apply?
- (4) Normative Model(s): What normative models underlie the standards of judicial review in the context of the relevant legal doctrines?
- (5) Philosophical Conceptions of Value: What philosophical conceptions of value underlie the normative models?

In his article on models for local government legitimacy,³⁶ Michelman illustrated the form and operation of this organizational structure. He examined three different factual situations: a local government's condemnation of private property for resale to a private developer in exchange for the developer's dedication of parking slots to the local government, the delegation of land use control authority to neighborhood groups³⁷ and attempts by local governments to zone out people living together in nontraditional living arrangements.³⁸ These factual circumstances were considered under three different doctrinal areas: public purpose, delegation of power to non-governmental groups and local government zoning with respect to nontraditional groups.

At the third level of analysis, Michelman suggested four different standards of judicial review for applying the legal doctrines involved: strictly procedural review,³⁹ rational basis substantive review,⁴⁰ per se categori-

tics and Values or What's Really Wrong with Rationality Review?, 13 CREIGHTON L. REV. 487 (1979); Michelman, *Political Markets*, *supra* note 19.

³⁵ The substance of this framework was derived from Michelman, *Political Markets*, *supra* note 19.

³⁶ *Id.*

³⁷ *Id.* at 147.

³⁸ *Id.* at 187-99.

³⁹ Under this form of review, a "court looks only to see that the [governmental agency] in reaching its decision has complied with the legislative procedures established by or under the constitution, statutes, and home rule charter if any, and that no specific rules against bribery, conflict of interest, or the like have been violated. The court [must] refuse[] to

cal substantive review⁴¹ and primary purpose or weighing-of-benefits substantive review.⁴² These forms of judicial review were themselves considered under two different normative models:⁴³ a public choice model and a public interest model.⁴⁴ He contends:

In the economic or public choice model, all substantive values or ends are regarded as strictly private and subjective. The legislature is conceived as a market-like arena in which votes instead of money are the medium of exchange. The rule of majority rule arises strictly in the guise of a technical device for prudently controlling the transaction costs of individualistic exchanges.⁴⁵

Thus, the public choice model, according to Michelman, is “an economic ideal conception of politics as a vehicle for private self-maximization.”⁴⁶ The public-interest model, on the other hand, according to Michelman,

depends at bottom on a belief in the reality -- or at least the possibility -- of public or objective values and ends of human action. In this public-interest model the legislature is regarded as a forum for identifying or defining, and acting towards those ends. The process is one of mutual search through joint deliberation, relying on the use of reason supposed to have persuasive force. Majority rule is experienced as the natural way of taking action as and for a group.⁴⁷

reexamine the merits or even the bare substantive rationality of the [governmental agency's] judgment.” *Id.* at 160.

⁴⁰ Under this form of review, a “court ascertains whether a [governmental agency], acting in good faith, could rationally conclude that the total (net) benefits to municipal citizens from this particular [governmental action] will exceed their total (net) tax and other costs.” *Id.* at 160-61.

⁴¹ Under this form of review, a court ascertains whether the governmental action violates a specific, applicable prohibition, such as the prohibition against the use of the city's taxing power in direct aid of a particular private enterprise, in the hypothetical situation posed. *Id.* at 161.

⁴² Under this form of review, a court ascertains whether the private benefits to the entrepreneur clearly predominate over the public benefits. *Id.*

⁴³ He defines a normative model of government as “a general conception of how governmental institutions ideally must be supposed to work in order to satisfy the conditions of a theory of moral justification for such institutions.” *Id.* at 147.

⁴⁴ *Id.* at 148.

⁴⁵ *Id.* See generally IAIN MCLEAN, PUBLIC CHOICE: AN INTRODUCTION (1987); CHARLES K. ROWLEY, DEMOCRACY AND PUBLIC CHOICE: ESSAYS IN HONOR OF GORDON TULLOCK (1987).

⁴⁶ Michelman, *Political Markets*, *supra* note 19, at 187. For a critique of this view of public choice theory, see Miller, *The Operation of Democratic Institutions*, 49 PUB. ADMIN. REV. 511 (1990) (arguing that public choice theory rejects a pluralist, “sum of the preferences” approach and is instead consistent with a liberal democratic ideal in the New Deal tradition).

⁴⁷ Michelman, *Political Markets*, *supra* note 19, at 149.

Thus, he considers the public interest model as “a noneconomic ideal conception of politics as a vehicle for community self-determination.”⁴⁸

At the fifth level of analysis, Michelman views these models as reflective of two different philosophical traditions: the public choice model is linked to a Hobbesian or Lockean notion of freedom:

The general idea is that values, so-called, are taken to be nothing but individually held, arbitrary and inexplicable preferences (the subjectivist element) having no objective significance apart from what individuals are actually found to be choosing to do under the conditions that confront them (the behaviorist element); from which it seems to follow that there can be no objective good apart from allowing for the maximum feasible satisfaction of private preference as revealed through actual choice -- or, in other words, through “willingness to pay.” The resulting allocation of resources to their highest-paying employments is the state known to economics as efficiency.⁴⁹

In contrast, the public interest model, he contends, traces its roots to a tradition in Western thought associated with Kant, Rousseau and Aristotle, that “conceives individual freedom in such a way that its attainment depends on the possibility of values that are communal and objective -- jointly recognized by members of a group and determinable through reasoned interchange among them[;]. . . freedom is the state of giving the law to oneself.”⁵⁰

Applying the two normative models to the reasoning and outcome of cases in the public purpose and delegation doctrinal areas of law, he concludes that public choice theory really consists of two variants: the “market failure” and “big bribe” variants. Under the big bribe variant, “we buy civil peace for the price of abiding by whatever the duly constituted legislative process presents us with.”⁵¹ Under the market failure variant, we accept governmental intrusion into the market when there is

unusual difficulty (high transaction costs) in striking a bargain or in organizing all who would have to agree to participate in a transaction in order that its potential, mutual benefits may be reaped. Such difficulty can result when the number of persons who stand to benefit from a costly undertaking is large, and none of these potential beneficiaries can produce the benefit for themselves without also making it available to others (whom it will not be possible to exclude, at any cost or at feasible cost).⁵²

Michelman demonstrates that a crucial element in the market failure variant is the theory of coalitions, (of log-rolling), whereby vote-trading occurs between legislators, because “[w]ithout that theory there can be no

⁴⁸ *Id.* at 187.

⁴⁹ *Id.* at 152-53.

⁵⁰ *Id.* at 150.

⁵¹ *Id.* at 162.

⁵² *Id.* at 155.

credible assurance that each individual will, over the long run, derive net advantage to his or her own, private interests from the majoritarian state.”⁵³ But, alas,

in recognizing the crucial significance of log-rolling, one sees that the market failure account depends not only on a conception of self-interested legislative traders -- or of legislative traders self-interestedly responsible to self-interested constituents -- but also on a conception of legislative output as not a series of discrete, separately intelligible and appraisable enactments, but rather a continuous unitary network of compromise -- of implicit multilateral trade so complex as to be almost certainly opaque and indecipherable to any outside observer.⁵⁴

Michelman thereby concludes that there is no room for substantive review under the market failure theory, whereas the “big bribe” theory supports only procedural review. Accordingly, court cases in these areas can be consistently explained by the public choice model only by restricting judicial review to a strictly procedural analysis: The governmental action must be upheld “as long as the [action] can be seen to have originated in a duly constituted legislative process in which there was the clearest of opportunities for log-rolling by or on behalf of a well-defined, “vested” interest ... which stood to be directly harmed by the ... enactment...”⁵⁵ The public choice model would thus render illegitimate any form of substantive review of governmental action, whether deferential or activist, and regardless whether such review is undertaken pursuant to the public purpose or delegation doctrines, general welfare provisions (police power), or the due process or equal protection clauses.⁵⁶

Applying the public interest model to the decisions of cases in these areas, Michelman suggests that courts would “not have available any tightly structured logic or formula for deducing when a legislature has switched out of the ideal role ... [of searching] for the right or best answer.”⁵⁷ “They would have to consult their own educated understanding of the values broadly shared in their society -- including the rate and direction of evolution of values -- in order to make judgments about whether given legislative products fairly reflect an effort to realize those values or trajectories ... [s]ometimes ... remitted to ... hunches.”⁵⁸

Methodologically, he contends, courts applying the public interest model must be particularly attuned to the motivational importance of context:

[T]he definition of aims through politics is an ethical process, and one which treats the individual as the ultimate object of ethical concern[, and

⁵³ *Id.* at 173.

⁵⁴ *Id.* at 175-76.

⁵⁵ *Id.* at 177.

⁵⁶ *Id.* at 177.

⁵⁷ *Id.* at 178.

⁵⁸ *Id.* at 178.

such insistence means that when individuals act politically, when they act as citizens, they . . . act on behalf of and with regard to one another, as well as themselves, as persons worthy of a full and equal measure of respect.⁵⁹

Thus, for example, citizens are more likely to perceive themselves as acting in a true civic forum, rather than a battlefield or market, if the decisionmaking group is not a narrow segment or a limited group.⁶⁰

V. THE UNSTATED ASSUMPTIONS OF CONVENTIONAL LEGAL ANALYSIS OF LOCAL GOVERNMENT AUTHORITY

Michelman's criticisms of the public choice model as an underlying normative framework for judicial decisionmaking in the public purpose, delegation and single-family zoning areas is persuasive. The crux of the matter is that substantive judicial review of local government action in these areas persists, and that the public choice model simply does not treat such review as legitimate.

A corollary of Michelman's analysis is that market failure in particular is not a potential underlying explanation for the outcomes in these areas, because, as he notes,

the market failure account depends not only on a conception of self-interested legislative traders -- or of legislative traders self-interestedly responsible to self-interested constituents -- but also on a conception of legislative output as not a series of discrete, separately intelligible and appraisable enactments, but rather [as] a continuous unitary network of compromise -- of implicit multilateral trade so complex as to be almost certainly opaque and indecipherable to any outside observer.⁶¹

Accordingly, as Michelman shows, only strictly procedural review, whereby a

court looks only to see that the [governmental agency] in reaching its decision has complied with the legislative procedures established by or under the constitution, statutes, and home rule charter if any, and that no specific rules against bribery, conflict of interest, or the like have been violated . . . [and does not] reexamine the merits or even the bare substantive rationality of the [governmental agency's] judgment⁶²

can legitimately take place under public choice theory.

Michelman's contention that public interest theory grounds judicial review in these areas is not as persuasive as his arguments about public choice theory. He does not convincingly show that the public interest model is a more likely underlying normative framework; he only shows that the public choice model is not substantiated by courts' actions in this area.

⁵⁹ *Id.* at 185.

⁶⁰ *Id.* at 186.

⁶¹ *Id.* at 175-76.

⁶² *Id.* at 160.

Perhaps the most telling difficulty with the public interest model is the lack of any authoritative sources of substantive content against which judges can evaluate legislative decisionmaking. As he points out,⁶³ “In the end, whether in the economic conception or [in the public interest conception,] the judge has to fall back on an educated sense of how people think, feel, or want.”⁶⁴

⁶³ *Id.* at 199.

⁶⁴ An alternative analytical model would not completely reject the Public Choice theory criticized by Michelman, but would instead incorporate economic efficiency as merely one potential objective. Under such an enriched approach, the sources of substantive content for judicial review of governmental action would be more precisely identified as deriving from various common law, statutory and constitutional provisions, both state and federal, and further shaped by the particular cognitive constructs which judges bring to the interpretive task: their basic experiences, idealized cognitive models and elaborating metaphors. Such an approach would provide a broader, more inclusive foundation for judicial review, taking into account not only “public interest” concerns, but “public choice” concerns as well.

Cass Sunstein has urged a similar focus on specific textual references as particular sources of substantive content for judicial review. Cass R. Sunstein, *Interest Groups in American Public Law*, 38 STAN. L. REV. 29 (1985) (hereinafter Sunstein, *Interest Groups*); Cass R. Sunstein, *Naked Preferences and the Constitution*, 84 COLUM. L. REV. 1689 (1984) (hereinafter Sunstein, *Naked Preferences*); See also Cass R. Sunstein, *Beyond the Republican Revival*, 97 YALE L.J. 1539 (1988) (hereinafter Sunstein, *Beyond the Revival*) (elaborating on his thesis by urging that courts adopt a more activist standard of judicial review as a general constitutional requirement, not just when discrete values or interests are threatened). His project is to develop a theory of legislative action to deal with the evil of “naked preferences,” defined as those circumstances in which “the distribution of resources or opportunities to one group rather than another solely on the ground that those favored have exercised the raw political power to obtain what they want.” Sunstein, *Naked Preferences*, at 1689.

Sunstein assimilates political ordering according to “naked preferences” to market ordering and contrasts that with “a conception of a political process in which differential treatment is justified not by reference to raw political power, but to some public value that the differential treatment can be said to serve.” Sunstein, *Naked Preferences*, at 1694. The latter maps almost perfectly onto Michelman's conception of a public interest model of legislative decisionmaking. Accordingly, we can view Sunstein's intellectual project as including the elaboration of forms of judicial review, as well as other, structural changes, that are necessary in order to assure a public-regarding legislative process. Sunstein, *Interest Groups*, at 63-64. This might supply the necessary content to standards of judicial review which seek to operate pursuant to the public interest model, a shortcoming left open in Michelman's *Political Markets* piece.

In summary, Sunstein seems to reject the proposition that values for informing public interest analysis are exogenous to the political process. Instead, he “envisions the Constitution and the courts imposing a deliberative model on Congress, through judicial review of the discussions and interactions between legislators, protection of legislators' independence from constituents, and stimulation of greater access of different groups to legislative deliberations, all with the purpose of ensuring that representatives can and will express their view of legislation in a thoughtful manner.” Michael A. Fitts, *The Vices of Virtue: A Political Party Perspective on Civic Virtue Reforms of the Legislative Process*, 136 U. PA. L. REV. 1567, 1589 (1988). Sunstein finds justification for such review in a “modern re-

The “Thayerite objection” remains, however: “[T]he tendency of a common and easy resort to [judicial review is] to dwarf the political capacity of the people, and to deaden its sense of moral responsibility” that comes from “fighting the question out in the ordinary way.”⁶⁵ That is, formulation of an acceptable theory of legislative action nevertheless requires that we formulate an acceptable role and theory for judicial review of such action.

In another article elaborating upon a modern republican theory of legislative action⁶⁶ Michelman acknowledges that judicial review of such action cannot be “strictly procedural review,” a point he also made in his *Political Markets* article.⁶⁷ In that article he concludes that:

The Court helps protect the republican state -- that is, the citizens politically engaged -- from lapsing into a politics of self-denial. It challenges

publican” understanding that “politics has a deliberative or transformative dimension(; its function is to select values, to implement “preferences about preferences,” or to provide opportunities for preference formation rather than simply to implement existing desires.” Sunstein, *Beyond the Revival*, at 1545.

⁶⁵ Michelman, *Law's Republic*, *supra* note 34, at 1525 n.127, quoting JAMES BRADLEY THAYER, JOHN MARSHALL 106-07 (1901) (1988).

⁶⁶ The modern republican understanding includes four central principles, each informing and defining the others:

(1) The first principle is deliberation in politics, made possible by what is sometimes described as “civic virtue.” In the deliberative process, private interests are relevant inputs into politics; but they are not taken as pre-political and exogenous and are instead the object of critical study.

(2) The second principle is the equality of political actors, embodied in a desire to eliminate sharp disparities in political participation or influence among individuals or social groups. Economic equality may, but need not, accompany political equality; in this sense, as in others, the political process has a degree of autonomy from the private sphere.

(3) The third principle is universalism, exemplified by the notion of a common good, and made possible by “practical reason.” The republican commitment to universalism, or agreement as a regulative ideal, takes the form of a belief in the possibility of settling at least some normative disputes with substantively right answers.

(4) The fourth and final principle is citizenship, manifesting itself in broadly guaranteed rights of participation. Those rights are designed both to control representative behavior and to afford an opportunity to exercise and inculcate certain political virtues. Citizenship often occurs in nominally private spheres, but its primary importance is in governmental processes. Sunstein, *Beyond the Revival*, *supra* note 64, at 1541-42.

The modern republican conception includes a theory of rights that “understands most rights as either the preconditions for or the outcome of an undistorted deliberative process.” *Id.* at 1551. Thus, liberty of expression and conscience, the right to vote (*Id.*) and freedom of religion (*Id.* at 1555) are basic preconditions for deliberation, the first principle of modern republican thought.

⁶⁷ Michelman, *Political Markets*, *supra* note 19, at 177; Michelman, *Law's Republic*, *supra* note 34, at 1525 (similarly distancing himself from “process-based” judicial review criticized by Laurence Tribe in Laurence H. Tribe, *The Puzzling Persistence of Process-based Constitutional Theories*, 89 *YALE L.J.* 1063 (1983), and “representation-reinforcing” judicial review suggested by John Hart Ely in JOHN HART ELY, *DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW* (1980)).

'the people's self-enclosing tendency to assume their own moral completion as they now are and thus to deny to themselves the plurality on which their capacity for transformative self-renewal depends.⁶⁸

The path he uses to get there begins with a philosophical conception of American constitutionalism which rests “on two premises regarding political freedom: first, that the American people are politically free inasmuch as they are governed by themselves collectively, and, second, that the American people are politically free inasmuch as they are governed by laws and not [people].”⁶⁹ Judicial review, Michelman thereby contends, is legitimate because it (merely) seeks to reinforce the (assumed) underlying basic tenets of American constitutionalism.

The unsettling fact remains, however, that it is a *judicial* determination of whether legislation is both “our” and “law.” He responds to that concern, at least in part, by saying that

[j]udges perhaps enjoy a situational advantage over the people at large in listening for voices from the margins. Judges are perhaps better situated to conduct a sympathetic inquiry into how, if at all, the readings of history upon which those voices base their complaint can count as interpretations of that history -- interpretations which, however re-collective or even transformative [--] remain true to that history's informing commitment to the pursuit of political freedom through jurisgenerative politics.⁷⁰

Having thereby overcome the Thayerite objection, at least tentatively, we confront the deeper “Cartesian Anxiety”: “the sense that we are caught between [dominating] objectivism, the belief that there are or must be some fixed, permanent constraints to which we can appeal and which are secure and stable[,] and [nihilistic] relativism, the message that there are no constraints except those we accept.”⁷¹ In other words, having formulated at least a tentative justification for judicial review, we must consider formulation of standards under which that review will be conducted.

In confronting the Cartesian Anxiety, Michelman has made some powerful arguments that the activity of judicial review of governmental

⁶⁸ Michelman, *Law's Republic*, *supra* note 34, at 1532. He thereby arrives at approximately the same place as Sunstein regarding the role of judicial review. *See supra* note 64.

⁶⁹ Michelman, *Law's Republic*, *supra* note 34, at 1499-1500.

⁷⁰ *Id.* at 1537.

⁷¹ Michelman, *The Supreme Court 1985 Term, Foreword: Traces of Self-Government*, 100 HARV. L. REV. 1, 24 n. 110 (1986) (quoting RICHARD J. BERNSTEIN, *BEYOND OBJECTIVISM AND RELATIVISM: SCIENCE, HERMENEUTICS, AND PRAXIS* 16, 18-19 (1983)). *See also* Winter, *Transcendental Nonsense*, *supra* note 8, at 1127 (“In its starkest form, it is the tension between foundationalism and nihilism[:]. Either there is some support for our being, a fixed foundation for our knowledge, or we cannot escape the forces of darkness that envelop us with madness, with intellectual and moral chaos.” quoting RICHARD J. BERNSTEIN, *BEYOND OBJECTIVISM AND RELATIVISM: SCIENCE, HERMENEUTICS, AND PRAXIS* 18 (1983)).

action is consistent with modern republican theory.⁷² He explains that judges themselves engage in dialogue, among multiple-judge panels and among courts. To that we might add the continuing dialogue among courts, executives, legislatures and voters expressing their preferences through elections, initiatives and referenda. But even Michelman's attempt at justifying "civic virtue"-reinforcing judicial review falls short of an explanation of what sources the court can use to test whether "civic virtue" has been achieved by legislative processes in light of any given substantive legislative outcome. As Feldman⁷³ points out, the statement that "valid" kinds of legislative action are based on public value is tautological: in essence, it justifies any government action as long as it is premised on more than the exercise of raw political power.⁷⁴ Feldman argues instead that preferences are socially constructed not only through the political process, but through the judicial process as well.⁷⁵ He suggests:

The Court's practice of constitutional adjudication is a conversation with the rest of society that requires the Court simultaneously to search for and to create meaning.⁷⁶

...

In Kuhnian terminology, we all experience and perceive reality through various paradigms--world views--and those paradigms consist of structures that are socially constructed or created.⁷⁷

The same can be said about judges. And, unlike Michelman's conceptions, standards of judicial review are themselves the judges' paradigms, or world views, through which they perceive reality. That reality may be articulated in terms of Michelman's conventional analytical framework, consisting of the philosophical foundations of law, normative models, legal doctrine and the facts, but it is nevertheless a product of that more elusive process we know as cognition.

⁷² Michelman, *The Supreme Court 1985 Term, Foreword: Traces of Self-Government*, 100 HARV. L. REV. 1, 74-77 (1986).

⁷³ Stephen M. Feldman, *Essay: Exposing Sunstein's Naked Preferences*, 1989 DUKE L.J. 1335, 1338 (1989) (hereinafter Feldman, *Essay*).

⁷⁴ See Sunstein, *Naked Preferences*, *supra* note 64, at 1694.

⁷⁵ Feldman, *Essay*, *supra* note 73, at 1335.

⁷⁶ *Id.* at 1357.

⁷⁷ *Id.* at 1341 citing THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS*, 10-13 (2d ed. 1970). See also THOMAS KUHN, *THE ESSENTIAL TENSION*, *SELECTED STUDIES IN SCIENTIFIC TRADITION AND CHANGE* (1977) (Kuhn's later thoughts on the nature of change); I. BERNARD COHEN, *REVOLUTION IN SCIENCE* (1985) (regarding the nature of changes in science generally); I. BERNARD COHEN, *THE NEWTONIAN REVOLUTION* (1980) (illustrations of the transformation of scientific ideas).

VI. A COGNITIVE CRITIQUE OF CONVENTIONAL LEGAL ANALYSIS

A. MICHELMAN'S APPROACH AS METAPHOR: "PYRAMID" THINKING

The analytical model Michelman presupposes may be expressed in cognitive terms as a triangle or as its three-dimensional cousin, a pyramid. The pyramid metaphoric representation of the Michelman analytical framework entails successively more abstract expressions of principles as one moves from the bottom to the top of the pyramid. Thus, the base of the pyramid is the level of individual fact patterns, unruly and incoherent in their myriad variety, doctrine exists at the next "higher" "level of abstraction" up the pyramid, then standards of judicial review, followed by normative models and, toward the top of the pyramid, philosophical conceptions of value.

This cognitive reconstruction of the Michelman approach reveals two powerful features of the Michelman model. The first feature is an implicit hierarchy of control. Each higher level of abstraction has a determinative effect on all the levels below it: doctrine determines the facts, the standard of judicial review determines the applicable doctrine, normative models determine the standard of judicial review and philosophical conceptions of value determine the normative models.⁷⁸ This feature of the Michelman analytical model may be a manifestation of the perhaps more familiar principle that our analytical models define our reality: we see what we expect to see. Thus, the doctrines we use to resolve disputes determine what facts are relevant, the standards of judicial review determine what doctrines we perceive as controlling, and so on up the analytical pyramid.

The second feature of the Michelman model viewed through a cognitive lens, however, is the perhaps more sinister normative claim that the higher up on the pyramid a concept appears, the "better" it is. Thus, for example, since the doctrinal level is "higher," the implicit normative claim is that it is a "better" way of dealing with reality than by confronting the naked chaos of its myriad factual variations. This implies that we should not experience chaotic reality directly, but that we should do so only indirectly, through multiple lenses. In other words, we are encouraged to assume that reality is orderly, if only we can see it the right way.⁷⁹

⁷⁸ For a description of this process as "synchronic," in contrast to a "diachronic," or narrative, rhetorical style, see Carol M. Rose, *Property as Storytelling: Perspectives from Game Theory, Narrative Theory, Feminist Theory*, 2 YALE J. OF L. & THE HUMANITIES 37, 37-39 (1990).

⁷⁹ There is an insidious prescriptive dimension to the pyramid: [T]he version of reality that has for the most part prevailed in the entire culture gives us internal scripts about how to argue and, indeed, how to know. The dominant culture has established certain criteria for theories, for legal arguments, for scientific proofs -- that is, for authoritative discourse. Thus, the very ground rules for disputing which version of reality should prevail belong to the world view that has been dominant in the past.

B. THE ORDER WE SEE IS THE ORDER WE IMPOSE

Michelman's way of seeing reality, however is just one way. Cognitive analysis allows us to critically examine that perspective. Cognitive analysis allows us to name Michelman's approach as a cognitive model and thereby establish an essence to which we can relate.⁸⁰ Moreover, cognitive analysis allows us to describe one feature of Michelman's cognitive model as incorporating the proposition that reality is orderly as its starting point. We can then proceed to consider whether reality is indeed orderly, or whether the order we see is instead the order we impose.

“Chaos” theory is a field of learning that touches on that assumption. Sometimes known as nonequilibrium theory or transformation theory, chaos theory “presents a view of the processes of change in which instability, disorder, and unpredictability serve as central features in the development of new forms of organization and complexity.”⁸¹ It attempts to explain how apparently random systems -- whether natural phenomena such as dripping faucets,⁸² roulette wheels,⁸³ or weather patterns⁸⁴ or social phenomena,⁸⁵ such as methods of public administration⁸⁶ or market economies⁸⁷ -- may suddenly and unpredictably experience dramatic transformations.

Chaos theory suggests that we should not assume that the world exists in patterns of order with intermittent moments of chaos, but that we should assume instead that the world is normally chaotic, with moments of order. Viewed in Chaos terms, legal decisionmaking--the resolution of social disputes through authoritative means, whether by courts, legislatures or popular votes--is not an orderly process, but a chaotic one. Chaos theory therefore suggests that legal decisions are *not* arrived at through the application of doctrine to facts, that doctrine is *not* determined by standards of judicial review, that such standards are *not* determined by normative

Martha Minow, *The Supreme Court, 1986 Term--Foreword: Justice Engendered*, 100 HARV. L. REV. 10, 65 (1987).

⁸⁰ “[O]ur acts of language are actions in the world, not just in our minds.” JAMES BOYD WHITE, *JUSTICE AS TRANSLATION: AN ESSAY IN CULTURAL AND LEGAL CRITICISM* ix (1990).

⁸¹ L. Douglas Kiel, *Nonequilibrium Theory and its Implications for Public Administration*, 49 PUB. ADMIN. REV. 544 (1989).

⁸² ROBERT SHAW, *THE DRIPPING FAUCET AS A MODEL CHAOTIC SYSTEM* (1984).

⁸³ THOMAS A. BASS, *THE EUDAEMONIC PIE* (1985).

⁸⁴ JAMES GLEICK, *CHAOS, MAKING A NEW SCIENCE* (1987); JOHN BRIGGS & F. DAVID PEAT, *TURBULENT MIRROR, AN ILLUSTRATED GUIDE TO CHAOS THEORY AND THE SCIENCE OF WHOLENESS* (1989).

⁸⁵ Loye & Eisler, *Chaos and Transformation: Implications of Nonequilibrium Theory for Social Science and Society*, 32 BEHAV. SCI. 53 (1987).

⁸⁶ Kiel, *Nonequilibrium Theory and its Implications for Public Administration*, 49 PUB. ADMIN. REV. 544 (1990).

⁸⁷ John L.R. Proops, *Organization and Dissipation in Economic Systems*, 6 J. SOC. & BIOLOGICAL STRUCTURES 353 (1983).

models, and that such models are *not* determined by philosophical conceptions of value. Instead, Chaos theory suggests that decisionmaking activity is chaotic, that legal decisions are the manifestations of dynamic systems--courts, legislatures or popular voting--whose outcomes are *not* orderly, but chaotic. "Law," in Chaos terms, therefore, is both a product and a manifestation of society, a society that is chaotic, not determined.⁸⁸

Chaos theory is therefore useful for us in two ways. First, it provides a counterpoint to Michelman's cognitive model: Michelman posits a reality that is orderly, with intermittent disorder; Chaos theory posits a reality that is disorderly, with intermittent periods of order. Second, at a "meta-cognitive model" level, Chaos theory serves not just as an alternative cognitive model to Michelman's approach, but also as a model for critical examination of *any* cognitive model. Cognitive models may thus be viewed as simply alternative ways to implement contested human ends. We need not settle on one model for all places and for all time; there are different ways to see the world. In this conceptualization, Chaos theory serves as a cognitive construct that is not just an explanation of behavior, but a possible vocabulary for describing available strategies.

The vocabulary of Chaos theory as a meta-cognitive model seeks to identify the underlying purposes served by any cognitive model. These purposes may include a need for a sense of control, a need for predictability, a need for narrowing disputes and shaping the terrain for the exercise of power and a need for an understanding of meaning and location. Cognitive models which may not look much like Michelman's approach at all may thus serve many of the same purposes.

VII. STORIES AS ALTERNATIVE STRATEGIES FOR DEFINING LOCAL GOVERNMENT AUTHORITY

A. THE TELLING POINT OF STORIES

A story is structured chaos.⁸⁹ It provides a narrative foundation from which the observer, (whether listener, reader or viewer), can imagine several possible outcomes. A crucial difference between conventional analytical

⁸⁸ Michelman acknowledged that his analytical framework did not explain all that courts do in the areas he studied. Courts do not, for example, follow one or another standard of judicial review, symbolizing one or another normative principle, and so on up the pyramid. Instead, courts seem at times to follow one, at times another principle, and at times some principle not included in the pyramid. In this fundamental way, Michelman's analytical framework simply demonstrates that it does not work as a completely satisfactory explanation of the very activity it purports to depict.

⁸⁹ For a classic examination of the intersection between law and storytelling, see, e.g., Richard Delgado, *Storytelling for Oppositionists and Others: A Plea for Narrative*, 87 MICH. L. REV. 2411 (1989). See also Michelle Oberman, *Your Work Will Be Your Most Faithful Mistress: Thoughts on Work-Life Balance Occasioned by the Loss of Professor Jane Larson*, 28 WIS. J.L. GENDER & SOC'Y 181 (2013) (telling a story about the corrosive effect on women and society of the having it all and work-life balance narratives).

discourse and the storytelling form of discourse is that the listener of a story is in large part responsible for what stories have to say.⁹⁰ In contrast, conventional analysis directs the observer toward certain pre-determined, and largely undebatable, content.⁹¹ Stories invite the listener to create a content for the story that is unique to the observer. Conventional forms of discourse, by contrast, control in large part what the observer derives from the experience of reading, listening or participating in the communication involved.

The activities of local government in relation to its citizens, to the state or federal legislature, or to the courts may be viewed as stories. Such stories provide the framework for analysis; they do not contain their own explanations. Alternative interpretations are accommodated; it is possible to derive different, yet acceptable, meanings.

*B. EXAMPLES OF ISSUES OF LOCAL GOVERNMENT POWER
CONSIDERED AS STORIES*

We can apply story rhetoric to the areas of local government authority examined by Michelman. The first situation involves a local government's condemnation of private property for resale to a private developer in exchange for the developer's dedication of parking slots to the government⁹². The story rhetoric would go something like this:

Centerville has experienced a substantial increase in automobile traffic in its downtown areas in the past two years. The consensus among the city council members is that additional parking is needed to accommodate the increased number of vehicles. Additional parking space could be created through construction of a downtown parking structure, but no public or private entity seems inclined to undertake such a project. The city has therefore decided to condemn a private piece of property in Centerville and resell the parcel to a private developer willing to construct a parking structure.

Downtown Parking Associates, an existing private parking concern, is willing to buy land in the downtown area for parking structure construction, but has been unable to obtain a piece of land to do so because all the downtown property owners want to hold off on selling because they believe downtown land will dramatically escalate in value in the years to come.

Downtown Parking has contacted the city and is willing to dedicate a couple of parking slots in a newly constructed structure for city use if the city will use its eminent domain power to condemn a suitable parcel of private land in the area.

⁹⁰ For an excellent sampling of Native American stories, presented in storytelling form, see HYEMEYOHSTS STORM, *SONG OF HEYOEHKAH* (1981); See also HYEMEYOHSTS STORM, *SEVEN ARROWS* (1972).

⁹¹ See Rose, *supra* note 79 (referring to conventional analysis as "synchronic," in contrast to a "diachronic," or narrative, rhetorical style).

⁹² See *supra* notes 36-38 and accompanying text.

The city council has decided to take up Downtown Parking Associates' offer, but has been challenged on its authority to do so.

At this juncture, alternative sub-plots to the story could be elaborated:

Disgruntled Competitor The challenger is another private developer who was willing to take the city's deal, but was unwilling to dedicate any parking slots for the city.

Environmentally-Concerned Citizen The challenger is a concerned citizen who wants no additional parking in the downtown area because it will generate additional pollution.

Traditionalist Citizen The challenger is a concerned citizen who thinks local governments should not be engaged in real estate development, but that such activities should be left in the private sector.

A major social value involved in defining the proper scope of local government authority in these circumstances is to assure public control over land use in the downtown area. Another purpose is to provide predictability in downtown land use so others in the downtown area will know how to structure their own uses. A third need is to assure that questions of land utilization will turn on a settled idea about whether there will be enough parking to accommodate the amount of traffic generated by a burgeoning downtown. Finally, a fourth need is to assure general confidence in the operations of the city council and the extent to which their decisions preserve the health, safety and welfare of the local residents.

Each of these purposes might be differently served by the implementation of the city's plan for a parking structure. Moreover, and significantly, the city might be considering different needs, depending on what parties are challenging its action. Unlike conventional doctrinal analysis, however, a story presentation does not foreclose the possibility that the dispute might be resolved in different ways, depending on who is challenging the proposal. Such a possibility is not part and parcel of the conventional understanding of dispute resolution through litigation. In essence, a story cognitive model frees the imagination to conceive of different ways to address a social conflict.

The second situation dealt with by Michelman involves definition of a local government's power to delegate land use control authority to neighborhood groups.⁹³ The story rhetoric would go something like this:

Several "anti-satellite disk" residents of a neighborhood are concerned about the proliferation of satellite disk antennas in their neighborhood. They would like to veto the installation of unsightly antennas in the area and have approached the city council with various proposals for achieving that objective.

Several sub-stories might arise:

Absolute Veto Under one proposal, the anti-satellite disk residents would have an absolute veto over any prospective disk an-

⁹³ *Id.*

tenna, without review by any governmental entity and without having to explain their reasons.

Reasonableness Review Under another proposal, the anti-satellite disk residents would only be empowered to exclude “unsightly” satellite disks from the neighborhood. Again, there would be no governmental review of their decisions and no standards by which to measure unsightliness.

Governmental Review Under a third proposal, the anti-satellite disk residents would have the power to preclude installation of “unsightly” satellite disks, but the city council would review all decisions of the residents as a matter of right to avoid circumstances in which “arbitrary” action was taken.

The purposes underlying the limitations on local government power to delegate authority to neighborhood groups include avoiding delegation of standardless discretion, assuring those affected by neighborhood group decisions have a proper opportunity to be heard and securing involvement of governmental officials whenever policies that transcend a neighborhood are involved.

Each of these purposes might be differently served depending on the criteria, procedures and opportunity for appeal to governmental authority that might be incorporated into the governmental delegation of power to neighborhood groups. Significantly, the concern for access to information of those who want to install satellite disks may invoke the need for more constrained delegation of authority to neighborhood groups. That interest is not tightly “localized.” In contrast, delegation of authority to neighborhood groups to control noise or the planting of pollen-sprouting plants that may generate allergic reactions throughout the neighborhood might call for different outcomes. Again, the story rhetoric allows for consideration of such variations.

The third situation dealt with by Michelman involves definition of a local government's power to zone out people living together in nontraditional living arrangements.⁹⁴ The story rhetoric would go something like this:

A city is concerned about preserving the character of its single-family residential (SFR) neighborhoods. Accordingly, it would like to define its SFR zones in such a way as to preserve the essence of the SFR zones. The essence of the SFR zone is viewed as the incorporation of only one “family unit” per structure. Thus, structures that would accommodate two families, such as duplexes, apartments or condominium structures, would not be consistent with the “single” family unit scheme.

There are at least two possible sub-stories available:

⁹⁴ *Id.*

Preservation of the Existing Intensity of Land Use The “single” families could be defined in terms of numbers of individuals occupying the same structure. Thus, rote numbers could be used to limit the size of a “family” for purposes of the ordinance.

Exclusion of Undesirable Living Arrangements “Single” families might be defined to exclude any living arrangements other than the most conventional, formally married, husband-and-wife-and-children grouping.

The purposes underlying the limitations on local government power to define “families” for purposes of single-family residence zones include preserving local authority to define local land uses, preventing local authorities from making distinctions that are not reasonably likely to achieve their objectives (preventing local jurisdictions from making arbitrary laws) and protecting vital individual interests from oppression by local governments.

Each of these purposes might be differently served depending on the criteria incorporated into a local government's definition of a “family.” Thus, if the definition of a family is so tightly circumscribed that only the storybook nuclear family would fit it, then only a very few groups of related individuals would qualify. The SFR zones would experience a constant turnover of “families” moving into and out of the area as the people involved reconstitute their living arrangements from a nuclear to an extended “family.” Although silly, this definition seems to have no other significant shortcomings.

Any definition meant to include inevitably excludes. Thus, the purpose behind the definition of a “family” is not necessarily objectionable. The real rub arises, of course, to the extent that the definition of the family utilizes characteristics that impinge on fundamental rights or use suspect criteria for classifying among acceptable and unacceptable residents. Thus, defining the “family” in stereotypical conventional terms would not seem objectionable, unless people are affected whose deeply held traditions involve extended, nontraditional living arrangements. In those circumstances, the sincerity of the beliefs in the alternative arrangements, the stability of the tradition in the culture, the extent to which the incorporation of such living arrangements would prejudice the limited use character of the SFR zones and the motivations behind the formulation of the definitions of “families” to exclude such living arrangements would certainly all come into play.

VII. Conclusion

A cognitive science approach frees the mind to conceive of different ways of resolving social conflicts. Liberated from the seemingly entrenched character of conventional understanding of legal doctrine, we can more easily debate about which cognitive pathways better serve contested human ends. This article suggests that conventional legal analysis is but one cognitive model for that work and that stories may prove valuable alternative tools.

COMMANDING CONSISTENTLY WITH SOVEREIGNTY:
THOMAS HOBBS'S NATURAL LAW THEORY OF
MORALITY AND CIVIL LAW

Kody W. Cooper*
Texas State University

ABSTRACT

In this essay, I seek to defend a reading of Hobbes's theory of civil law as properly a natural law theory. On the natural law account, for some ordinance to fully bind one with the force of civil law, in the focal sense of law, it must be both systemically and morally valid.¹ While, strictly speaking, it is anachronistic to apply the notions of systemic and moral validity to assess Hobbes's theory of civil law, I suggest that the distinction is helpful in making sense of Hobbes's theory. Hobbes's account of civil law qualifies as a natural law account because for him, a datum is systemically valid if it is sourced in the sovereign's command and it is morally valid if it does not contravene natural law. To defend this interpretation I first sketch an interpretation of Hobbes's theory of morality as a properly natural law theory. I argue that Hobbes's relation to the high water mark of the classical natural law tradition, as expressed in Thomistic thought, is different than is often supposed. I then proceed to consider his account of civil law and take up two interpretive puzzles in Hobbes that seem to obscure the distinction between systemic and moral validity: Hobbes's theories of authorization and the "mutual containment" of the natural law and positive law. I offer a solution to each interpretive puzzle that would maintain the distinction. On this basis, Hobbes's theory of civil law is vindicated as properly a natural law theory. I will begin by sketching Hobbes's natural law account of morality, which provides the foundation for his account of civil law.

* Lecturer, Department of Political Science, Texas State University. I am indebted to A.P. Martinich, J. Budziszewski, Devin Stauffer, and one anonymous reviewer at the *Journal* for comments on earlier drafts of this article.

¹ For discussion and defense of this thesis of natural law theory, see JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS, 26, 351-66 (2d ed. 2011); John Finnis, *Natural Law: The Classical Tradition*, in THE OXFORD HANDBOOK OF JURISPRUDENCE AND PHILOSOPHY OF LAW 20-23, 33-34 (Jules L. Coleman & Scott J. Shapiro eds., 2002); For Joseph Raz's distinction between systemic and moral validity, see JOSEPH RAZ, THE AUTHORITY OF LAW 146-59 (1979).

CONTENTS

I.	INTRODUCTION	166
II.	HOBBS' NATURAL LAW THEORY OF MORALITY.....	167
III.	HOBBS ON THE CIVIL LAW.....	179
IV.	TWO OBSCURITIES: AUTHORIZATION AND MUTUAL CONTAINMENT	182
V.	AUTHORIZATION, SOVEREIGNTY, AND INALIENABLE RIGHTS	185
VI.	THE MUTUAL CONTAINMENT THESIS AND THE SOVEREIGN RIGHT OF JUDGMENT.....	189
VII.	CONCLUSION	196

I. INTRODUCTION

Thomas Hobbes's theory of law is one of the most memorable in the history of legal and political thought. As the frontispiece to *Leviathan* vividly portrays, Hobbes constructs a theory of absolute, united sovereignty out of a mass of individual wills which make up the body politic. Hence, the frontispiece portrays an imposing crowned figure, brandishing sword and crozier, with a body made up of citizens. In the midst of the English Civil War, Hobbes's message seemed to be that, in order to secure the peace, the sovereign must have the kind of power over the body politic that an individual has over his or her bodily members—an absolute power to command motion proper to their capacities. And since law is the principal act by which the sovereign moves the body politic, it seems to follow that Hobbes's command theory of civil law should place him as a positivist in the history of legal and political thought. And yet, Hobbes has much to say about the natural law, a body of unwritten precepts that have an extra-human source that secures their legal character, apparently independent of the civil sovereign's will. When this textual fact about Hobbes's corpus is taken into account, it is not so clear how Hobbes should be understood in history of legal thinking. Norberto Bobbio writes:

Thomas Hobbes belongs, de facto, to the history of the natural law tradition. There is no text on the history of legal and political thought which does not mention and analyse his philosophy, as one of the typical expressions of natural law theory. On the other hand Hobbes belongs, de jure, to the history of legal positivism. His conception of the

law and the state is indeed a surprising anticipation of nineteenth-century positivist theories.²

Bobbio himself goes on to offer what seems to have become a fairly standard reading of Hobbes as a positivist in which Hobbes's statements about natural law are subsumed under a positivist interpretation of his conception of civil law.

In what follows I seek to challenge the positivist interpretation of Hobbes. I contend that Hobbes holds a peculiar natural law theory of morality that grounds his theory of civil law. Hobbes's theory of the human good and conception of the natural law precepts as sourced in the divine will provides a content-based criterion for judging the legal validity of sovereign acts. I recount a couple of key tensions or obscurities in Hobbes's thought regarding authorization and inalienable rights and defend a natural law interpretation of Hobbes's theory of sovereignty. I conclude by suggesting a better way to understand Hobbes's place in the history of legal philosophy.

II. HOBBS'S NATURAL LAW THEORY OF MORALITY

Hobbes's account of morality counts as a natural law theory because he retains two key notions that classical natural law theory considered requirements for something to count as a *natural law theory*: the human good, which is grounded in human nature, provides basic *reason(s)* for action and the norms or precepts that correspond to the human good have a *legal character*. The distinctiveness of Hobbes's natural law account of morality and civil law in the broader natural law tradition flows chiefly from his *thin theory of the good*. Let us consider classical natural law theory's thick theory of the good and the legal character of natural law precepts.

The core notion of classical natural law theory lies in those standards—principles, rules, or norms which give or purport to give direction in deliberation about what to do—of right judgment in matters of practice (conduct or action). We can speak of these standards as *natural* inasmuch as they are not the product of individual or collective choice and not subject to repeal—"however much they may be violated, defied, or ignored"—because mere individual or collective choice cannot change the kind of thing man is.³ And, we can speak of these standards as *lawful* inasmuch as they bind or ought to bind in one's deliberations about what to do. These rules, norms, or laws are rooted in the first principles of practical reason, which are fittingly described as those most *basic reasons for action* that direct us to the range of human goods. Since Thomas Aquinas is considered to be the gold standard of classical natural law theory, let us consider Aquinas's thick theory of the good in his presentation of natural law theory.

² NORBERTO BOBBIO, THOMAS HOBBS AND THE NATURAL LAW TRADITION 114 (Daniela Gobetti trans., 1993).

³ Finnis, *supra* note 1, at 1-2.

Aquinas's presentation of the thick theory of the human good proceeds according to what John Finnis fittingly calls a "metaphysical stratification" of human nature: (1) what we have in common with all substances, (2) what, more specifically, we have in common with other animals, and (3) what is peculiar to us as human beings.⁴ Hence, in Aquinas's formulation, the human goods include: preservation of one's substantial being, marriage and childrearing, friendship with others in society, and knowledge of the truth, including the truth about God. These goods make up the objective content of happiness or authentic human well-being because they are objectively knowable by all rightly reasoning persons. While Finnis's own presentation of human good differs in texture from Aquinas's, his account correctly re-presents the traditional natural law view that each human good provides a non-instrumental reason for action—the good or reason for acting is in some sense *basic*.⁵ Corresponding to these goods is the order of precepts of the natural law, i.e., the norms regarding preservation of human life, sex and education of children, shunning ignorance, and living peaceably with one's fellows. This is, in very short outline, classical natural law's *thick theory of the good*.

It is also a sketch of classical natural law theory's grounds for judging the moral validity of human positive law, since the flourishing of individuals and communities in their pursuits of the various forms of the human good is the standard guiding those who are charged with care of the whole community, when they deliberate about what to enact, decide, require, promote, etc. Since that which authorities have care over is a *communitas communitatum*, a community of communities, the authority's charge will be twofold. First, he must foster and protect the unity and well-being of the range of communities that enjoy intrinsic common goods, including the *communiones* of friendships, families, and religious believers. Second, he must foster and protect unity and well-being of the community at large. In other words, classical natural law theory held legislators are or ought to be guided by the common good.

Regarding the second requirement for something to qualify as a natural law theory, as Finnis points out, for Aquinas, the ultimate source of

⁴ FINNIS, AQUINAS 81 (1998). I agree with Finnis's characterization of Aquinas in this respect. But, as will be apparent in what follows, I differ from Finnis in that I would affirm Aquinas has a hierarchical account of the goods, and that the divine command is essential to the legal character of the natural law.

⁵ Finnis's own presentation of the human goods or the most basic reasons for action includes: bodily life and health, friendship, marriage, knowledge, skillful performance in work and play, harmony between one's inner and outer life, and harmony with the ultimate source of reality (Finnis, *Natural Law Theory and Limited Government? in NATURAL LAW, LIBERALISM, AND MORALITY: CONTEMPORARY ESSAYS* 4 (Robert P. George ed., 1996). Notably, Finnis has sought to present the goods and moral norms of classical natural law theory while accepting the fact/value dichotomy and denying a hierarchy of the goods prior to choice. Finnis has been criticized by more traditional natural law theorists for these moves. See, e.g., RUSSELL HITTINGER, *A CRITIQUE OF THE NEW NATURAL LAW THEORY* (1987).

reality enhances “both the content and the normativity” of the first principles.⁶ The norms of natural law have a *legal character*. How is that?

For Aquinas, the basic norms of natural law have the character of *law* because they meet the four necessary conditions for something to *be* law: (a) an ordinance of reason (b) for the common good (c) made by a proper authority and (d) promulgated.⁷ Aquinas believes natural law is *law* because he holds a vision of the universe—all of “nature,” including human nature—as *created* and *ordered* by a providential and loving God (doctrines that Aquinas believed were demonstrable by unaided reason in the science that we would today call philosophical theology). Human beings in particular are ordered toward a form of well-being available only to rational creatures. The beatitude available to man by his unaided powers is an end that specifies good and bad action. Good acts are those acts ordered to happiness and bad acts are those acts not ordered to happiness or flourishing. Those goods that are basic or the basic reasons for action ground precepts that, while not sufficient to secure one’s full-fledged flourishing, keep one from falling off the cliff in one’s moral life. For Aquinas, the precepts take on the character of *law* prior to human positive law, inasmuch as God—the being who has care of the common good of the whole universe—promulgates them or makes them known, in the very act of creating and ordering man with reason and will. Moreover, since Aquinas holds that law is properly the command of an authority, Aquinas maintains that the natural law is commanded in God’s act of creating nature.⁸

Suppose we take Aquinas’s theory to be the apex of classical natural law theory. On this understanding, *modern* natural law theory breaks from classical natural law theory in at least two ways: in its treatment of practical reasoning as essentially in the service of sub-rational passions and in its secular foundations. Hume stated the modern view most sharply when he claimed reason is and only can be a slave of the passions and in his skepticism of natural theology. But, on Finnis’s reading, the modern understanding of practical reason as enslaved to the passions is traceable to Thomas Hobbes.⁹ I call this understanding of practical reason the *impotent thesis*,

⁶ FINNIS, *supra* note 4, at 128, 308.

⁷ It is said that natural law “maximally has the character of law”—*lex naturalis maxime habet rationem legis*. This is the premise of an objection that Aquinas does not deny (THOMAS AQUINAS, *SUMMA THEOLOGIAE* I-II 90, 4, obj. 1) [hereinafter ST]. The focal case of law for Aquinas is, of course, the eternal law—yet, since natural law is not diverse from the eternal law, but rather a mode of the eternal law’s promulgation, Aquinas could affirm the premise, properly understood (*Id.* at 91.2, ad. 2).

⁸ *Id.* at 95.3; cf. *DE VERITATE* 17.3. In his intellectualist sense of command, Aquinas can be counted as a member of what Gerald Postema calls the “command tradition” of law (Gerald J. Postema, *Law as Command: The Model of Command in Modern Jurisprudence*, 11 *PHIL. ISSUES* 470-501 (2001)).

⁹ As Finnis puts it, Hobbes treats “our practical reasoning as all in the service of sub-rational passions such as fear of death, and desire to surpass others—motivations of the very kind identified by the classical natural tradition as in need of direction by our reasons’ grasp of more ultimate and better ends, of true and intrinsic goods, of really

because it claims that practical reason does not have the power to set its own goals. Indeed, the impotent thesis is the standard interpretation of Hobbes's theory of practical reason. Hence, standard interpretations of Hobbes's natural law theory tend to posit a universal desire, to which reason is instrumental. Given its strong textual basis in Hobbes's corpus, the universal desire typically posited is the desire for self-preservation—and this is supposed to secure the normativity of the laws of nature.

Moreover, the standard interpretation of Hobbes's natural law theory includes what we can broadly call the secularist thesis, the claim that God plays no substantive role in Hobbes's moral and political thought. This claim is defended on the basis of three sub-theses: the historical thesis, the concealment thesis, and the practical severability thesis. God plays no substantive role because Hobbes's contemporaries accused him of atheism and impiety, indicating actual textual meaning (the historical thesis),¹⁰ or because the bits of irony sprinkled throughout his texts indicate that his theistic and religious claims are so many genuflections to the religious authorities of his day (the concealment thesis)¹¹ or because even supposing Hobbes is a theist, God is irrelevant to his political philosophy (the practical severability thesis).¹² On the secularist view, Hobbes's laws of nature are mere "qualities" or "theorems" and do not attain the status of law until the erection of an absolute sovereign.¹³ While these features of the standard interpretation—the pure instrumentality of practical reason and secularism—have not gone unchallenged, they probably remain the conventional wisdom.

But, these two features of the standard interpretation of Hobbes's natural law theory—the impotent thesis and the secularist thesis—do not fit well with two principles Hobbes holds: first, the psychological diversity of persons and second, the eternal, immutable, and universal bindingness

intelligent reasons for action. Hobbes proclaims his contempt for the classical search for ultimate ends or intrinsic reasons for action. Accordingly there can be for him no question of finding the source of obligation and law in the kind of necessity which we identify when we notice that some specific means is required by and for the sake of some end which it would be unreasonable not to judge desirable and pursuit-worthy." (Finnis, *supra* note 1, at 5). For similar, influential narratives of Hobbes's place in the history of natural law theory, see LEO STRAUSS, *THE POLITICAL PHILOSOPHY OF HOBBS*, (Elsa M. Sinclair trans., 1952) and Michael Zuckert, *Do Natural Rights Derive from Natural Law?*, 3 HARV. J.L. & PUB. POL'Y 695-731 (1997).

¹⁰ See SAMUEL MINTZ, *THE HUNTING OF LEVIATHAN* 45 (1962); Quentin Skinner, *Hobbes's 'Leviathan'*, 7(2) HIST. J. 332 (1964).

¹¹ See STRAUSS, *supra* note 9; Edwin Curley, *I Durst Not Write So Boldly, or How to Read Hobbes' Theological-Political Treatise in HOBBS E SPINOZA: SCIENZA E POLITICA* (Daniela Bostrenghi ed., 1992).

¹² See GREGORY S. KAVKA, *HOBBSIAN MORAL AND POLITICAL THEORY* 362 (1986).

¹³ See also, PEREZ ZAGORIN, *HOBBS AND THE LAW OF NATURE* 51 (2009).

of the laws of nature, *in foro interno*.¹⁴ Call these the *psychological diversity principle* and the *bindingness principle*. Regarding the first, Hobbes observes a number of cases in which persons fail to desire self-preservation. Hobbes believed that people may be and often are willing to lay down their lives for the sake of personal honor or what Sharon Lloyd has called “transcendent interests.”¹⁵ Recognizing the force of this point, one might water down the putatively necessary desire for self-preservation to be a *predominant* desire in order to make it more psychologically fitting. But, this option is ruled out if we take seriously Hobbes’s second principle regarding the eternal, immutable, and *universal* bindingness of the laws of nature, because then the laws of nature would only bind usually or for the most part. They would not bind universally, since not everyone actually has the putatively universal desire. In short, “if [the laws of nature] are always to bind everyone *in foro interno*, their claim on us must either depend on no desires, or on a desire that no human can fail at any time to have.”¹⁶

Now, this may simply be another example in which Hobbes is simply “inconsistent and irreconcilable” with himself as Bramhall alleged was evident in a whole range of Hobbes’s doctrines.¹⁷ Or they may be instances in which Hobbes was, in his own words, “a forgetful blockhead.”¹⁸ But, Hobbes’s texts actually suggest another possibility, namely, that practical reason identifies bodily life and health as a—indeed, *the*—basic good or *reason for action*. Hobbes indicates as much when he lays down reason and cupidity as the two postulates of human nature. Hobbes lays down two postulates of human nature in the Dedicatory Epistle to *De Cive*: first, the postulate *cupiditatis naturalis*, whereby man demands private use of common things and second, the postulate *rationis naturalis*, which teaches man to avoid violent death or “fly contra-natural dissolution [*mortem violentam*]” as the greatest natural evil.¹⁹

While cupidity is the principle of covetousness in man—which, unchecked, leads to the widespread destruction and misery in the state of nature—the rational principle “teaches every man to fly a contre-naturall Dissolution, as the greatest mischief that can arrive to Nature.” It has appeared to some that Hobbes here identifies reason with the passion of

¹⁴ THOMAS HOBBS, *LEVIATHAN WITH SELECTED VARIANTS FROM THE LATIN EDITION OF 1688*, 15.18, 99; 26.40, 186 (Edwin Curley ed., 1994). Citations to *LEVIATHAN* [hereinafter L] are to chapter, paragraph, and page number in this edition.

¹⁵ On the psychological diversity of Hobbesian persons, including the Hobbesian agent’s concern for transcendent interests, see S.A. LLOYD, *MORALITY IN THE PHILOSOPHY OF THOMAS HOBBS: CASES IN THE LAW OF NATURE* 56-94 (2009).

¹⁶ *Id.* at 200.

¹⁷ John Bramhall, *The Catching of Leviathan or The Great Whale* in *THE WORKS OF THE MOST REVEREND FATHER IN GOD*, Vol. IV, 869 (1844).

¹⁸ THOMAS HOBBS, *ENGLISH WORKS VOL. IV* 387 (William Molesworth ed., 1840) [hereinafter EW].

¹⁹ EW II, viii; THOMAS HOBBS, *DE CIVE: THE LATIN VERSION* 75 (Howard Warrender ed. 1983).

fear.²⁰ Yet, I contend that the tenor of the passage is to *distinguish* between reason and desire sufficiently to indicate they are at cross-purpose in man.²¹ And this suggests that reason is not or need not be a slave to the passions. It also suggests that Hobbes is not a sheer value subjectivist.²² On this reading, as Bernard Gert has argued, the goal of practical reason, to fly contra-natural dissolution, is independent of the contingent desires of natural cupidity.²³ In other words, *life*, which Hobbes refers to as the *bonum maximum*, is the basic good in Hobbes's axiology.²⁴ In this way, I suggest that Hobbes's contrast with the classical natural law tradition lies not in the sheer instrumentality of practical reason, but in his *thin theory of the (objective) good*. On this basis, Hobbes holds a thin conception of the common good or peace as chiefly security. While I think Finnis correctly states the traditional natural law view in his claim that it is "unreasonable and unrealistic ... to treat survival as the sole basic reason for acting,"²⁵ I suggest that the thinness of Hobbes's notion of the good does not in itself discount his theory from being a *natural law theory*—but it does mark it off as novel in relation to the older tradition.²⁶

Such a reading saves both the psychological diversity principle and the bindingness principle because, while all persons may not actually take the

²⁰ See, e.g., STRAUSS, *supra* note 9.

²¹ I defend this interpretation at greater length in my *Reason and Desire After the Fall of Man: A Rereading of Hobbes's Two Postulates of Human Nature*, 107-29 HOBBS STUDIES XXVI (2013).

²² Stephen Darwall argues against subjectivist interpretations of Hobbes. See Stephen Darwall, *Normativity and Projection in Hobbes*, 109 PHIL. REV. 313-47 (2000). However, if my argument is correct, Darwall's "projectivist" solution mistakenly reads Hume back into Hobbes.

²³ In my view, Gert has given powerful textual evidence for this interpretation of Hobbesian reason as more than a merely reckoning or computing power. The power of Hobbesian practical reason includes the ability to set its own goals. See Bernard Gert, *Hobbes on Reason*, 82 PAC. PHIL. Q. 243-57 (2001).

²⁴ *De Homine* [hereinafter, DH] 11.6 in OPERA PHILOSOPHICA QUAE LATINE SCRIPSIT OMNIA: IN UNUM CORPUS NUNC PRIMUM COLLECTA STUDIO ET LABORE VOL. II 98 (William Molesworth ed., 1839-1845) [hereinafter OL]. While scholars have taken Hobbes's denial of a *summum bonum* as a rejection of any objective ends *simpliciter*, nothing in that passage actually requires that interpretation. Hobbes qualifies the claim to apply to felicity *in this life* and *as is spoken of the books of old moral philosophers* (L 11.1, 57; DH 11.15). Supposing Hobbes was a sincere Christian, he may only be saying perfect beatitude is not available in this life—and Aquinas explicitly affirms this as well (ST, I-II, 2.8).

²⁵ Finnis, *supra* note 1, at 27. Finnis makes this remark in criticism of H.L.A. Hart, who maintained that reasons for action besides survival were controversial.

²⁶ On this point I agree with Mark Murphy. See Mark Murphy, *Was Hobbes a Legal Positivist?*, 105 (4) ETHICS 846-73 (1995). However, I differ from Murphy in that I defend Hobbes's theory of the good as "reason-dependent" to use Joseph Raz's phrase—see JOSEPH RAZ, THE MORALITY OF FREEDOM 140-44 (1986) and I show how Hobbes understands the laws of nature to be truly *laws*, a point that Murphy claims Hobbes does not explain.

good of life as basic in their practical reasoning, *they rationally ought to*. We might say that this is the basic requirement of practical reasonableness. The laws of nature can then be understood as so many practical necessities that conduce to the basic good of life. Beyond the basic good of life, the subjectivity of the good will be more stated and thus Hobbesian felicity will be widely inclusive of many life-plans. But, the basic requirement of a reasonable life-plan will be taking the good of life as basic.

This interpretation may strike some as initially implausible. After all, doesn't Hobbes famously portray the state of nature as a place where everyone has a right to all things, i.e., a place devoid of moral requirements? Not exactly. When Hobbes defines the right of nature as "the liberty each man hath, to use his own power, as he will himself, for the preservation of his own nature; that is to say, of his own life," Hobbes does not lay down an absolute liberty, but a right tethered to the goal of securing the good of life.²⁷ Since Hobbes believed that duties always come with the means to fulfill them, the right of nature can be understood to be the necessary means to carry out the rational necessity to preserve one's life.

Still, doesn't Hobbes maintain that "whatsoever is the object of man's appetite or desire; that is it, which for his part he calls *good*"?²⁸ And doesn't Hobbes say that the term good is "ever used in relation to the person that useth them: there being nothing simply and absolutely so" and that there is no "common rule" of the good in the object itself?²⁹ Wouldn't this undermine any notion of an objective good in Hobbes's moral philosophy? While these passages have been taken as the basis for the subjectivist interpretation of Hobbes's metaethics of the good, I believe that a better interpretation can be offered that fits with Hobbes's psychological diversity and bindingness principles.³⁰

In the first quoted passage, Hobbes need not be taken to be claiming anything more than that commonly, folks in fact "calleth good" whatever they in fact desire. Hobbes's claim that nothing is simply or absolutely good need only imply that there is no actually good object that is not *in fact* being desired. I suggest that Hobbes is trying to express a point that Peter Geach later made along these lines that "there is no such thing as being just good or bad, there is only being a good or bad so and so."³¹ Geach made the point by reflecting on how we talk about the good. In the common way we speak about objects, good is an attributive adjective—it "sticks" to the noun that it modifies as in the proposition, "X is a good car." In contrast with the proposition "X is a red car," the former proposition does not split up into "X is good" and "X is a car." Whereas, one

²⁷ L, 14.1, 79.

²⁸ *Id.*, at 6.7, 28.

²⁹ *Id.*, at 6.7, 29.

³⁰ See DAVID GAUTHIER, *THE LOGIC OF LEVIATHAN* (1967); KAVKA, *supra* note 12; JEAN HAMPTON, *HOBBS AND THE SOCIAL CONTRACT TRADITION* (1986).

³¹ Peter Geach, *Good and Evil* reprinted in *THEORIES OF ETHICS* 64-73 (Phillipa Foot ed., 1967). Lloyd makes a similar argument. See LLOYD, *supra* note 15, at 83.

could see the that the distant object is red and one's color blind friend can see that it is a car, "there is no such possibility of ascertaining that a thing is a good car by pooling independent information that it is good and that it is a car." Similarly, Hobbes is saying that to claim that some object or state of affairs is *good* cannot be said simply independent of the object's relation to human desire. So when an agent says, "that car is good," the agent is typically saying something relative to his or her purposes (e.g.) it is good *for* transporting me to work or good *for* impressing the 4H babes, or good *for* off-roading, etc. Lloyd seems correct that Hobbes's peculiar way of speaking does not undermine the objectivity of the good, but makes it rather opaque.³² In short, Hobbes does not maintain that "good" is just synonymous with "desired by the agent." The objectivity of the good of life seems still available. While more work is needed to fill out this interpretation of Hobbes's metaethics of the good, the foregoing provides an outline of my defense of the claim that Hobbes retains the first essential feature of a properly *natural law theory*, namely that the good, rooted in human nature, provides basic reasons for action.

The second essential feature of a properly *natural law theory* is that the precepts of morality have a *legal* character. Hobbes offers a theological warrant for his claim that the practical necessities that conduce to the good of life are eternally, immutably, and universally binding *in foro interno* with the force of law. *Law* is properly the command of one who has the right to so command, and God, who enjoys that right in virtue of his omnipotence, commands the laws of nature by his rational word.³³ Hence, contrary to the practical severability thesis, I suggest that God does play an essential role in Hobbes's natural law theory, because the divine legislative pedigree secures the legal character of the laws of nature. Hobbes's theological and religious views are admittedly an ongoing flashpoint of debate in Hobbes scholarship.³⁴ Yet, in my view, the work of A.P. Martinich has recovered what is salvageable of the so-called Taylor-Warrender thesis and shown that there are good grounds to reject the secularist theses.³⁵ Hence, I suggest that we are warranted in supposing that "for the most part,

³² LLOYD, *supra* note 15, at 83.

³³ L, 15.51, 200, 31.3, 235, 31.7, 237.

³⁴ For a recent discussion of leading parties in the debate and a sophisticated argument for the skeptical reading of Hobbes, see Patricia Springborg, *Calvin and Hobbes: A Reply to Curley, Martinich, and Wright* 4 PHIL. READINGS 3-17 (2012); Springborg, *Hobbes's Challenge to Descartes, Bramhall, and Boyle: A Corporeal God* 20 BRIT. J. HIST. PHIL. 903-34 (2012).

³⁵ The secularist theses face a host of historical, textual, and hermeneutical difficulties. See generally A.P. MARTINICH, *THE TWO GODS OF LEVIATHAN* (1992); A.P. Martinich, *Interpretation and Hobbes's Political Philosophy*, 82 PAC. PHIL. Q. 309-31 (2001). For Martinich's assessment and reply to Springborg's case for the skeptical reading, see A.P. Martinich, *On Thomas Hobbes's English Calvinism: Necessity, Omnipotence, and Goodness* 4 PHIL. READINGS 18-30 (2012) and *Epicureanism and Calvinism in Hobbes's Philosophy: Consequences of Interpretation* PHIL. READINGS (forthcoming).

Hobbes meant what he said.”³⁶ I will thus proceed by taking seriously Hobbes’s claims that the legal character of the laws of nature is grounded in a conception of God as creator and orderer of man.³⁷

Yet, is God really necessary to secure the legal character of the precepts of morality, the laws of nature? Is not the power of reason to set its own goals sufficient to secure the legal character of the laws of nature, independent of whether they have a divine legislative pedigree? Bernard Gert is one example of the many sophisticated interpreters of Hobbes who have suggested as much.³⁸ But, how can the mere operation of practical reason impose a duty, on Hobbes’s terms? Hobbes’s account suggests that obligation is an essentially relational notion. For example, in a covenant, when Jones gives up his right to ϕ to Smith for some determinate time in exchange for something from Smith now, Jones forms a promissory obligation that he fulfills by not ϕ -ing or by Smith releasing him from his promise.³⁹ How, then—apart from any covenant (or a contract)—does one have a natural or rational duty? Hobbes is clear that autonomous practical reason cannot self-impose a duty: “Nor is it possible for any person to be bound to himself; because he that can bind, can release; and therefore he that is bound to himself only, is not bound.”⁴⁰ While Jones might say he is legislating a law for himself, such an act would not generate an obligation because Jones would be both promisor and promisee—and whatever he decides to do would satisfy the “obligation.”⁴¹ It follows that, for Hobbes at least, the *legal force* of the laws of nature in conscience, *prior* to contracts in the state of nature and the sovereign-making covenant, cannot be generated by the mere operation of individual practical reason.⁴²

Nor is the horizontal relationship between human agents the only obligation-generating relation that Hobbes recognizes. On the contrary, Hobbes explicitly avers that the right to rule and hence the power to bind springs “either from nature or from contract.”⁴³ Hence, Hobbes holds that natural duty binds by the vertical relation that “is not by nature taken away.”⁴⁴ Now, Gert agrees with the dominant view in Hobbes scholarship that God plays no substantive role in Hobbes’s moral philosophy.⁴⁵ But the

³⁶ MARTINICH, *supra* note 35, at 16.

³⁷ I agree with Martinich that Hobbes is trying to wed the new science and materialism with the older theistic natural law picture—whether they actually compatible or can be so wed is of course another question.

³⁸ Gert, *supra* note 23, at 78, 82.

³⁹ *L*, 14.11, 82.

⁴⁰ *Id.* at 26.6, 174.

⁴¹ Matthew B. O’Brien makes this point in the context of his incisive criticism of Kant. See O’Brien, *Practical Necessity: A Study in Ethics, Law, and Human Action* (2011) (Ph.D. dissertation, The University of Texas at Austin).

⁴² As various readers of Hobbes have attempted to argue. See, e.g., Darwall, *supra* note 22.

⁴³ EW II, 206.

⁴⁴ *Id.*

⁴⁵ Gert, *supra* note 23, at 70.

foregoing suggests that, on Hobbes's own terms, Hobbes's dictates of reason can only be recommendatory and not moral *laws*, absent a divine legislative pedigree. For, Hobbes's claim that reason *teaches* us to fly death sounds more akin to the recommendation of counsel than the demand of command.⁴⁶ But the recommendatory force of counsel falls somewhat short of the exceptionless, binding force of *law, in foro interno*.⁴⁷ And we have seen that for Hobbes law cannot be generated through self-legislation. What is needed to secure the *legal* character that Hobbes attributes to reason's dictate is an interpretation of Hobbes's natural law theory in which God's command plays an essential role. I maintain that Hobbes's conception of God as creator and orderer of man secures the legal force of the laws of nature.

Hobbes's affirmation that the world is created—a claim Hobbes repeatedly makes in the first two parts of *Leviathan*—is grounded by his natural and revealed theology.⁴⁸ I shall briefly sketch his ground in unaided reason. Hobbes offers a number of arguments for God's existence and the various arguments he gives have similarities to Aquinas's five ways.⁴⁹ By the "light of Nature" we can know that the first attribute of God is *existence*.⁵⁰ *Existence* is the first of the divine attributes, because something cannot *be* a subject of an attribute, except insofar as it exists. But, if God exists, then there is no perfection that can be denied of him—indeed, all perfections are *maximally* so in God.⁵¹ Therefore, the judgment that God exists entails that God is maximally or, as Hobbes prefers to put it, *irresistibly* powerful, because power is a perfection. God is omnipotent. And if God is omnipotent he has complete power over all of nature. But God could not have maximal power over nature unless he *created* it, because if he didn't create it, then its existence would not depend on God's power, and then he would not have complete power over nature. Hence, if God exists, God created and ordered the world. But man is a part of the world.

⁴⁶ Cf. Hobbes's argument against Cardinal Bellarmine where he links teaching and counsel. The ecclesiastical prerogative is to *teach* not to *command* such that their teachings are "not laws, but wholesome counsels" (EW III, 490).

⁴⁷ L, 15, 99.

⁴⁸ See, e.g., *Id.* at 31.15-16, 239.

⁴⁹ See Robert Arp, *The 'Quinque Viae' of Thomas Hobbes*, 16 HIST. PHIL. Q. 367-94 (1999).

⁵⁰ L, 31.14, 239. Cf. ST I, 4.1, ad. 3.

⁵¹ The worship we do God "proceeds from our duty, and is directed according to our capacity, by those rules of honour that reason dictateth" and honoring "consisteth in the inward thought, and opinion of the power, and goodness of another: and therefore to honour God, is to think as highly of his power and goodness, as is possible" (L 31.8, 237; 31.13, 239). "Moreover in attributes which signify greatness or power, those which signify some finite or limited thing are not signs at all of an honouring mind" (EW II, 214). "He therefore who would not ascribe any other titles to God than what reason commands, must use such as are either *negative*, as *infinite, eternal, incomprehensible*, etc; or *superlative* as *most good, most great, most powerful*, etc." (EW II, 216). Compare Aquinas's "fourth way," the argument from the gradation of things ST I, 2.3.

Therefore, the order evident in man toward the good of life, and the order of the dictates of right reason to secure this good, is an order willed by a being with the right to will it.⁵² It has, in other words, the character of law. *Pace* proponents of the practical severability thesis, these claims cannot be explained merely as Hobbes's attempt to "multiply men's motives for following" the laws of nature or merely to show they are "compatible with Christianity" in order to persuade his Christian audience.⁵³ Rather, they are essential to secure the legal character of the laws of nature. Another way to see the point is to consider God as the author of reason and speech.

Significantly, Hobbes affirms that God is the "first author" of speech.⁵⁴ Hobbes's natural theology provides a warrant for this claim. Since we cannot deny a perfection to God, we must affirm that God is maximally rational. But speech is the mark of rationality. Since God is the first cause of all of nature, including man, and man is a rational animal, God must be the first author of speech. Hence, Hobbes's natural theology grounds his claim that God's legislation secures the legal character of the laws of nature, since God's "rational word" corresponds to the hearing of "right reason."⁵⁵ It is by the rational faculty that we judge life to be the basic good and that we rightly reason about the necessary means to preservation.

In short, Hobbes has one foot in the older natural law tradition, inasmuch as he insists that the basic norms of morality enjoy the status of *law* by their divine legislative pedigree. Hobbes's principal break with that tradition is his thin theory of the good.⁵⁶

If this account is correct, then a very different picture of Hobbes's account of civil law arises than the one we are accustomed to. There would be solid grounds to say, on Hobbesian terms, that the fundamental judgment that life is to be preserved and right reason's deductions of the neces-

⁵² Hobbes sees an analogy between the order in man-made artifacts and the order in that "most excellent work" of God, man: "For what is the *heart*, but a *spring*; and the *nerves*, but so many *strings*; and the *joints*, but so many *wheels*, giving motion to the whole body, such as was intended by the artificer?" (L, Introduction.1, 3). The evident order in man toward life is a particular instance of the "admirable order" we observe in the "visible things of the world," from which it is a valid inference that there is a cause of that order, which men call God (L, 11.25, 62).

⁵³ See LLOYD, *supra* note 15, at 104; Gert, *supra* note 23, at 70, 82.

⁵⁴ L, 4.1, 16. In my view, Philip Pettit's otherwise illuminating study of Hobbes's theory of language too quickly dismisses the significance of this thesis for Hobbes's thought. See PHILLIP PETTIT, *MADE WITH WORDS* 26 n.1 (2008).

⁵⁵ L, 31.3, 235.

⁵⁶ It is of course an ongoing metaethical dispute whether law and morality needs a foundation beyond the natural human good. On my interpretation, Hobbes believed such a ground was required. For an influential argument that "law talk" in ethics cannot be sustained without a theistic foundation, see G.E.M. Anscombe, *Modern Moral Philosophy*, 33 *PHIL.* 124 (1958). For one influential statement of non-theistic moral realism, see Michael S. Moore, *Good Without God*, in *NATURAL LAW, LIBERALISM, AND MORALITY: CONTEMPORARY ESSAYS* (Robert P. George ed., 1996).

sary means thereto—including fundamental law of nature’s directive to seek peace—bind with the force of law prior to the erection of the sovereign.

It is necessary to note that Hobbes’s thin theory of the good leads him to formulate a peculiar distinction between *in foro interno* and *in foro externo* bindingness of the laws of nature:

The laws of nature oblige *in foro interno*; that is to say, they bind to a desire they should take place: but *in foro externo*; that is, to the putting them in act, not always. For he that should be modest, and tractable, and perform all he promises, in such time, and place, where no man else should do so, should but make himself a prey to others, and procure his own certain ruin, contrary to the ground of all laws of nature, which tend to nature’s preservation.⁵⁷

The *in foro interno-in foro externo* distinction is intended to hold in balance two claims that are in apparent tension: first, that all of the laws of nature are eternally, immutably, and universally binding, and second, that morality is good in the modes of the *pulchrum*, the *jucundum*, and the *utile* (good in the promise, in the effect, and in the means), i.e., that the laws of nature conduce to the good of life.⁵⁸ Because the basic rational necessity to preserve one’s life is considered as prior, the catalogue of laws of nature, while always binding *in foro interno*, can fail to bind *in foro externo*, if one finds oneself in a mass-scale worst case scenario like the state of nature. They can fail when *the sufficient security condition* fails. That is, the laws of nature can fail to bind *in foro externo* because there can fail to be an entity with sufficient power to sanction noncompliance with the laws of nature, in which case obedience *in foro externo* would make one prey to others and contravene the more basic rational necessity to pursue life. When seen in this light, the right to all things in the state of nature is instrumental to the duty to preserve oneself, since duties come with the means to fulfill them.⁵⁹ The laying down of the right to all things in the sovereign-making covenant is done *for the sake* of peace—in contrast with war, in which the good of life is severely threatened—which is another way of saying that the sovereign is created in order to meet the sufficient security condition of the laws of nature obtaining *in foro externo* force. Hence, the moral validity of civil law must turn on whether its demands and permissions are congruent with one’s basic duty to preserve one’s life. On these grounds, I defend the following interpretation of Hobbes’s theory of civil law.

The basic good of life, which is made into the common good of peace through the sovereign-making covenant, amounts to a content-based criterion for the moral validity of civil law. Commands of sovereigns must conform to the natural law to achieve the status of law. In other words, com-

⁵⁷ L, 15.36, 99.

⁵⁸ *Id.* at 7.8, 29.

⁵⁹ *Id.* at 18.8, 113.

mands of sovereigns must be genuinely ordered to peace or the common good of security. But, as Perez Zagorin has recently pointed out, the natural law interpretation of Hobbes's moral theory faces the challenge of according with Hobbes's positivist-sounding claims about the civil law.⁶⁰

III. HOBBS ON THE CIVIL LAW

When Hobbes discusses civil law in relation to natural law, he makes a number of positivist-sounding remarks that suggests the ultimate source of law is the will of the commonwealth or sovereign. The suggestion is apparent in Hobbes's definition of civil law:

*Civil Law, is to every subject, those rules, which the commonwealth hath commanded him (by word, writing, or other sufficient sign of the will) to make use of, for the distinction of right, and wrong, that is to say, of what is contrary, and what is not contrary to the rule.*⁶¹

Hobbes's theory of civil law has been taken to be an expression of legal positivism when legal positivism is understood principally to mean this: the existence of a law depends on its pedigree, irrespective of its merits or content.⁶² As indicated in the quoted passage, the pedigree lies in *having been commanded*. This has appeared to some to be an early version of what latter-day positivist theory refers to as the "sources thesis." According to the sources thesis, the truth of statement "Legally, Jones ought to \emptyset " or "It is the law that Jones ought to \emptyset " depends on "an appropriate social fact specifiable without resort to moral argument."⁶³ Hence to know whether some action is legally demanded or permitted requires one to advert to a relevant social fact—in Hobbes's case the relevant social fact would consist in the will of the sovereign commander, as expressed in word, writing, or some sufficient declaration of will. The sources thesis entails some version of the "separability thesis."⁶⁴ That is, since the legal status of any datum depends solely on its pedigree, its status as a law does

⁶⁰ ZAGORIN, *supra* note 13.

⁶¹ L, 26.3, 173.

⁶² For interpretations of Hobbes as a legal positivist, see J.W.N. WATKINS, *HOBBS'S SYSTEM OF IDEAS* 114 (2d ed. 1973); KAVKA *supra* note 12, at 248-50; HAMPTON, *supra* note 30, at 107- 10; M.M. Goldsmith, *Hobbes on Law, in THE CAMBRIDGE COMPANION TO HOBBS* 275 ff (1996). For a recent assessment, see ZAGORIN, *supra* note 13, at 2-3, 49-54. For critiques of positivist readings of Hobbes, see Murphy, *supra* note 26; David Dyzenhaus, *Hobbes and the Legitimacy of Law*, 20 *LAW & PHIL.* 461-98 (2001). I differ from the latter critics in my interpretation of Hobbes as a natural law theorist. That is, I don't believe that they have explained the legal character of the laws of nature prior to human law, on Hobbes's terms.

⁶³ RAZ, *supra* note 1, at 65.

⁶⁴ Sometimes this is formulated as "there is no necessary connection between law and morality" and alongside the sources thesis constitutes the "core commitments of positivism" (Jules L. Coleman & Brian Leiter, *Legal Positivism*, in *A COMPANION TO PHILOSOPHY OF LAW AND LEGAL THEORY* 241 (Dennis Patterson ed., 1996).

not turn on its moral content, because, like its pedigree, the content of the law is a matter of social fact.

Of course, legal positivists today largely would not defend Hobbes's (supposed) version of legal positivism. H.L.A. Hart, the master of contemporary positivists, famously rejected older versions of positivism, including Austin's gunman account, Kelsen's *grundnorm* account, and Holmes's predictive account of law as over-simplified theories that failed in various ways to accurately describe our experience of law. Yet, Hart retained the core thesis of legal positivism, namely, that the *existence* and *content* of law is merely a matter of social fact, or proper pedigree. Hart himself identified Hobbes as a member of the positivist pedigree in the history of legal philosophy.⁶⁵ But, is it true that the existence and content of law are merely a matter of pedigree, independently of its moral content, according to Hobbes?

Hobbes declares that *law*, in its proper acceptance, *binds*. Law, says Hobbes, "determineth, and bindeth."⁶⁶ It is not that law merely *claims* to obligate but that, as indicated by the definition of law in general, it is of the very nature of law to obligate. So Hobbes also says that "law in general, is not counsel, but command; nor a command of any man to any man; but only of him, whose Command is addressed to one formerly obliged to obey him."⁶⁷ Hence, for something to *be* positive law—in the focal sense of *law*—is for it to *obligate*: it binds one, in conscience, to act or forbear. In Hobbes's lingo, it binds both *in foro interno* (in conscience) and *in foro externo* (putting it into outward act).

So then, since, by definition, the sovereign's command to his subjects is civil law, and since law of its very nature binds, it might appear that Hobbes is a legal positivist—the existence and content of the law are known by reference to the sovereign's will, and the law binds. That would mean that any command of the sovereign would attain the status of *civil law*—and hence, bind one to act or forbear—regardless of its content. But the plot thickens when one considers that Hobbes also indicates that part of the reason *law* binds, properly speaking, is because it provides one with a sufficient reason for action:

Commanding, which is that speech by which we signify to another our appetite or desire to have any thing done, or left undone, for reason contained in the will itself: for it is not properly said, *Sic volo, sic jubeo*, without that other clause, *Stet pro ratione voluntas*: and when the command is a sufficient reason to move us to the action, then is that command called a *law*.⁶⁸

In this passage, Hobbes seems to indicate that a command is only properly called law when it provides one with a sufficient reason for ac-

⁶⁵ H. L. A. HART, THE CONCEPT OF LAW 187-195 (1961).

⁶⁶ L, 14.3, 79.

⁶⁷ *Id.* at 26.2, 173.

⁶⁸ EW IV, 74-5.

tion.⁶⁹ In other words, the fact of some edict's *having been commanded* is not sufficient for it to attain the status of civil law (which, of its very nature, binds in conscience).

As we saw above, the basic reason for action in Hobbes's scheme is life or self-preservation.⁷⁰ Because God, who rightfully governs nature by his omnipotence, orders man toward life—as is evident in the order of our bodily parts and right reason—the duty to preserve one's life binds with the force of law, for rational (reasonable) actors. It follows that any command that would require one to destroy one's life could not provide a sufficient reason for one to act. I suggest this is how we ought to understand Hobbes's claim that there are certain actions that a man can never be bound by the sovereign's command to do: “not to defend a [his] own body,” “to kill, wound, or maim himself,” “not to resist those that assault him,” “to abstain from the use of food, air, medicine, or any other thing, without which he cannot live,” to accuse himself or self-incriminate without the assurance of pardon, or to serve as a soldier.⁷¹

Hobbes's catalogue of inalienable rights is explicable in terms of the basic rational necessity to preserve one's life, because one always acts blamelessly when does that which is necessary to perform one's duty. Sometimes this has been put as the “ought implies can” principle. Disobeying a command not to defend or nourish oneself is always done with right, because such acts would likely cause one's death. But, any command of the sovereign that attains the status of civil law is not disobeyed with right. It follows that the good of life is a content-based limitation that bounds the set of the sovereign's commands that attain the status of civil *law*. A command to do or forbear acts that would destroy one's life would lack moral validity for the addressee. In the older natural law tradition, such commands were called “perversions of law.”

Aquinas deploys this locution, *perversitas legis*, when considering the nature of a tyrannical command and is willing to use phrases like “unjust law” and “corruption of law” for edicts that contravene natural law.⁷² In locutions like “unjust law” “corrupt law” and “perverse law,” the adjectives ‘unjust’, ‘corrupt’, and ‘perverse’ modify *law*. Aquinas's way of speaking suggests he would agree that some iniquitous enactment could have systemic validity. Hence, to summarize Aquinas's position by merely quoting the “unjust law is not a law” dictum, without this nuance, would be

⁶⁹Cf., MURPHY, *supra*, note 26, at 853.

⁷⁰ Deploying Razian terminology to explain Hobbes's view, Susanne Sreedhar refers to the basic good of self-preservation as a “non-excludable first-order reason for action”(SUSANNE SREEDHAR, HOBBS ON RESISTANCE: DEFYING THE LEVATHIAN 108-31 (2010)). While I would not go so far as Sreedhar in attributing Raz's sophisticated account of law to Hobbes, I believe Sreedhar's discussion helps illuminate in various ways what Hobbes was trying to do—and my interpretation is in some ways compatible with her account.

⁷¹ L, 21.11-13, 141-2.

⁷² ST I-II, 92.1, ad. 4; cf. ST I-II, 93.3, ad. 2, ST I-II, 95.2, ST I-II, 94.6, ad. 3.

misleading. Aquinas can accept a partial truth in the legal positivist's sources and separability theses.⁷³ Any legislative enactment with systemic validity by that very fact has in some measure the character of law. However, inasmuch as it contravenes natural law, it lacks moral validity and hence fails to be law, in the focal sense of law—it will, in other words, lack the full character of law. The natural law theorist thus can accept the distinction between systemic and moral validity and still insist that mere systemic validity does not suffice to tell us whether some edict is the peripheral or focal case of *law*—i.e., whether it binds in conscience *as law*.

Admittedly, Hobbes himself does not use the locution “perversion of law.” But, the point is that, like Aquinas, Hobbes seems to recognize that commands of a sovereign can lack moral validity. Moreover, since both Aquinas and Hobbes hold that all moral value is ultimately rooted in God's will—abstracting here from the important question regarding the degree to which Hobbes's apparently more voluntarist conception breaks from Aquinas's more intellectualist conception—the claim that civil law must be morally valid to have the binding character of law ultimately means that it must be congruent with God's will. Hobbes breaks from Aquinas in formulating a thin theory of the good as the criterion of moral validity.

These initial reflections have moved too quickly, however, because we have so far abstracted from passages in which Hobbes obfuscates the distinction between moral and systemic validity. The success of my interpretation will require an account of these obfuscations. The distinction between systemic and moral validity in Hobbes is obscured for at least two reasons. The first reason is due to Hobbes's collapsing of a law's pedigree and its justice under the will of the sovereign, a collapse that he grounds in his theory of authorization. The second is his claim that the law of nature and the civil law are of “equal extent”—the claim that Kavka has aptly termed the “mutual containment thesis.”⁷⁴ Let us consider these two reasons for the obscurity.

IV. TWO OBSCURITIES: AUTHORIZATION AND MUTUAL CONTAINMENT

While Hobbes requires that law provide one with a sufficient reason for action—a claim that we have suggested introduces an inchoate distinction between systemic and moral validity in Hobbes's account of civil law—Hobbes obscures this distinction in his account of the justice of the sovereign's command.

The obscurity is evident if we look at Hobbes's answer to the following two questions, that can be asked of any putative law: (i) *Is this enactment systemically valid?* and (ii) *Is this enactment just?* Hobbes says that

⁷³ For a discussion, see Finnis, *The Truth of Legal Positivism in THE AUTONOMY OF LAW* 195-214 (Robert P. George ed., 1999).

⁷⁴ KAVKA, *supra* note 12, at 248.

both questions can be answered by knowing the answer to just one question: (iii) *Was the datum commanded by the sovereign?* Hence, Hobbes collapses the answers to questions (i) and (ii) into the facticity of the sovereign's command. Hobbes does this because, in his view, the sovereign's command is always the source of positive law and is always just.

Hobbes offers a few different arguments as to why the sovereign's command is never "unjust." The arguments revolve around Hobbes's understanding of the sovereign-making covenant.⁷⁵ His principal ground in *Leviathan* is his theory of authorization in the sovereign-making covenant. In the covenantal formula, a person says, "I authorize and give up my right of governing myself and authorize all the sovereign's actions."⁷⁶ Since a covenanter authorizes the sovereign to do whatever he (it) will do to him as a subject, the sovereign cannot be accused of injustice.⁷⁷ It follows that the sovereign's command cannot be *unjust*—"no law can be unjust."⁷⁸ Since Hobbes suggests that the answer to questions (i) and (ii) depends solely on the sovereign will, the distinction between systemic and moral validity is obscured—but it remains to be seen whether it is destroyed. Consider the second and related way that Hobbes obfuscates the distinction.

The second reason that the distinction between moral and systemic validity is obscured is Hobbes's "mutual containment thesis." Hobbes puts it this way: "The law of nature, and the civil law, contain each other, and are of equal extent."⁷⁹ This claim apparently has two parts. The first part, the containment of the civil law in the natural law, seems straightforward. The laws of nature direct men to lay down their right to all things, erect a sovereign power, and perform their covenant made. The meaning of the second part of the thesis is not immediately clear. How is the natural law contained in the civil law? Clearly, the civil law is supposed to enforce the dictates of reason with a power sufficient to sanction noncompliance. But, does it mean something more than that? An available and influential interpretation is that of Sharon Lloyd.

According to Lloyd, the mutual containment thesis is indicative of what she calls Hobbes's "self-effacing" natural law theory.⁸⁰ The natural law *itself* directs agents to authorize an unassailable judge to determine what the law—including both civil and natural law—*is*. Such an interpretation of Hobbes as a "practical legal positivist," for Lloyd, means that there is no legitimate perspective from which to criticize the sovereign. In other

⁷⁵ One example of such an argument is based on Hobbes's technical definition of justice as non-violation of covenant. Because the sovereign is not party to the sovereign-making covenant—subjects covenant *between themselves* to grant their rights to a sovereign—the sovereign is not party to the covenant, and so cannot be accused of injustice.

⁷⁶ L, 17.13, 109.

⁷⁷ *Id.* at 18.6, 112-113.

⁷⁸ *Id.* at 30.20, 229.

⁷⁹ *Id.* at 26.8, 174.

⁸⁰ S.A. Lloyd, *Hobbes's Self-Effacing Natural Law Theory*, 82 PAC. PHIL. Q. 285-308 (2001), at n.27. Cf., LLOYD, *supra* note 15, at 263-94.

words, natural law is contained in the civil law because an interpretation of the natural law is morally valid just in virtue of its having been interpreted by the sovereign and posited.

In short, Hobbes obscures the distinction between systemic and moral validity in two ways: the reduction of the justice of law to the sovereign's command by the theory of authorization and the mutual containment thesis, which Lloyd maintains is Hobbes's way of saying that natural law itself requires the erection of a sovereign with absolute authority to judge the meaning of natural law.

The first obscurity is nothing new. Hobbes's earliest critics pointed out the apparent contradiction between his sovereign-making formula and the right to resist the sovereign in self-defence, including Robert Filmer, George Lawson, and Bishop Bramhall.⁸¹ The apparent contradiction is manifest in our foregoing considerations. By the sovereign-making formula, a person gives up his rights of self governance to authorize the sovereign. From this formula, Hobbes infers that the sovereign can never do his (its) subjects injustice, since the subject authorized all his actions. But we also saw how Hobbes maintains that the right to preserve one's life is inalienable. Accordingly, we saw that Hobbes lists a number of acts that the sovereign can never bind a subject to do that would entail his self-destruction. And that which one cannot be bound by law to do is done with right, just as one cannot be taken to authorize an absolutely unlimited sovereign, if that entails the transfer of an inalienable right.⁸² So apparently, the subject does and does not authorize an absolute sovereign; the subject does and does not act justly when he or she disobeys the sovereign's command to act (or forbear) in a way that would destroy one's life.

Again, these obscurities may be instances of self-irreconcilability or forgetful blockheadedness. But, there may be solutions available on Hobbes's own terms.

Hobbes's apparently contradictory theory of authorization and inalienable rights seems to obscure the distinction between systemic and moral validity of law. But if it can be shown that the absolute justice of the sovereign's commands is compatible with an inalienable right to self-defense, then, upon that ground, the morally-systemically valid distinction would hold. And, on that ground, our thesis that Hobbes's account of civil law is properly a natural law account can be vindicated. Regarding the mutual containment thesis, the challenge would be to show how this thesis is compatible with the claim that civil law can fail to be morally valid and how the thesis does not entail a practical legal positivism. Let us first turn to see how Hobbes's theory of authorization, sovereignty, and inalienable rights might be clarified.

⁸¹ See the selections of their critiques of Hobbes in *LEVIATHAN*, Appendices A-C (A.P. Martinich ed., 2002).

⁸² L, 14.8, 82.

V. AUTHORIZATION, SOVEREIGNTY, AND INALIENABLE RIGHTS

How can we reconcile Hobbes's claim that the sovereign-making covenant authorizes an unlimited sovereign with his claims that one always retains those inalienable right necessary to preserve one's life? I will argue that the sovereign acts unlimitedly *inasmuch as he (or it) is sovereign*. That is to say, when the sovereign acts *as sovereign*, his (its) command is sufficient to make something into civil law. But the *inasmuch as* qualification turns out to import the content-based limitation that the natural law places on what can achieve the status of law for the addressee of a command.

When Hobbes is discussing those rights that are inalienable, he remarks that, when a person makes a covenant, he must always be understood to act under the aspect of the good: "the object is some *good to himself*."⁸³ But this claim is compatible with Hobbes's distinction between apparent and actual goods—and we have already seen that the basic, actual good for Hobbes is life.⁸⁴ Hobbes's axiology is best understood as a thin theory of the good objectively knowable by unaided reason. Hobbes appears to be saying that someone who enters into a covenant can be presumed to meet the minimum condition of practical reasonableness, that they take the good of life as basic. Hence, anyone who enters the sovereign-making covenant is presumed to take life as good. So Hobbes says in the same passage: "the motive, and end for which this renouncing, and transferring of right is introduced, is nothing else but the security of a man's person, in his life, and in the means of so preserving life, as not to be weary of it."⁸⁵ From these points, Hobbes concludes that someone who performs a covenantal act should never be taken to forfeit the end for which the covenant was made:

And therefore if a man by words, or other signs, seem to despoil himself of the end, for which those signs were intended; he is not to be understood as if he meant it, or that it was his will.⁸⁶

Earlier we saw Hobbes's covenantal formula included a person's authorization of all the sovereign's acts. If we take Hobbes at his word, then such an authorization cannot be a sign that would despoil the covenanter of the end for which he covenants, namely, the security of his person and the means to preserve his life. I suggest that the covenantal formula authorizes a person to command a set of ordinances, O1, *inasmuch as they act as sovereign*. The set O1, is distinct from the set of *all possible ordinances*, O2. My claim is that if, and only if, the command is of the set O1 can it achieve the status of civil *law* for the addressee. How can a sovereign fail to act *as a sovereign*?

⁸³ *Id.* at 14.8, 82.

⁸⁴ DH, 11.6.

⁸⁵ L, 14.8, 82.

⁸⁶ *Id.* at 14.8, 82.

Hobbes clearly recognizes that the person or persons with sovereignty act in ways that cannot be considered actions *as sovereign*. The sovereign (whether a man or an assembly) bears “two persons”—his own natural person, and the person of the commonwealth. Hence, the monarch “hath the person not only of the commonwealth, but of also of a man.”⁸⁷ When a sovereign acts “as a man” or in his “natural capacity” his (its) acts are not understood as representative of his subjects. Thus, Hobbes distinguishes between public ministers that are empowered by the sovereign to administer the realm and servants of a monarch who serve him in his “natural capacity.” When a sovereign orders his ministers, his act is essentially different from when he orders his private servants. In the latter case, he does not act *as bearer* of the commonwealth and hence does not act with the authority of the sovereign. When the sovereign does not speak *as the sovereign representative*, his words do not attain the status of civil law.

Still, the example of a sovereign giving orders to servants, stewards, chamberlains, and the like does not get to the heart of the difficulty we are interested in. The controversy lies precisely in potential scenarios in which the sovereign, in his public capacity, commands one to perform acts destructive of one’s life. I want to claim that such commands fail to achieve the status of civil law for the addressee of the command because when the commander so acts, he (it) is not acting *as sovereign*.

Consider Hobbes’s claim that when public ministers act in the name of the sovereign,

Every subject is so far obliged to obedience, as the ordinances he shall make, and the commands he shall give be in the king’s name, and *not inconsistent with his sovereign power*.⁸⁸

The passage indicates that there are potential ordinances that would be inconsistent with the sovereign power—ordinances that if commanded would not oblige subjects. But what is the criterion for inconsistency? There are of course a number of rights that inhere in the sovereign power.⁸⁹ But the office of the sovereign includes not *only* rights and powers needed to duly execute the office—it also includes the end to which those powers are ordained, i.e., the end for which the office was created:

The office of the sovereign, (be it a monarch, or an assembly) consisteth in the end, for which he was trusted with the sovereign power, namely the procuration of the safety of the people; to which he is obliged by the law of nature, and to render an account thereof to God, the author of that law, and to none but him.⁹⁰

Indeed, Hobbes indicated the telic nature of the sovereign power before either the covenantal formula or any of the essential rights of sovereignty:

⁸⁷ *Id.* at 23.2, 156.

⁸⁸ *Id.* at 22.3, 156.

⁸⁹ *See generally, id.* at 17.

⁹⁰ *Id.* at 30.1, 219.

Commanding Consistently with Sovereignty

The final cause, end, or design of men, (who naturally love liberty, and dominion over others,) in the introduction of that restraint upon themselves, (in which we see them live in commonwealths,) is the foresight of their own preservation, and of a more contented life thereby; that is to say, of getting themselves out from that miserable condition of war.⁹¹

The criterion of inconsistency is nothing other than that for which the covenant was formed. Importantly, this statement of the end of the sovereign-making covenant comes in the same chapter that Hobbes gives us the covenantal formula. Hobbes does not include the *telos* in the actual formula because he has already stated it, and, as before, we can presume that covenanters are practically reasonable. Just as any covenanter is presumed to be taking the good of life as basic, any person who participates in the authorization of the sovereign by the covenantal formula is presumed to be quitting the condition of war, for the sake of security. That is, covenanters authorize the sovereign to secure the peace, or to procure the common good.

Hobbes's view of the common good of peace, in contrast with the Thomistic natural law tradition, is thin: it aims somewhere between (i) the mere absence of civil strife and (ii) agreement of citizens on important matters. Hobbesian covenanters do not aim at a thicker notion of peace higher on the scale of the unity of peace. They do not aspire to (iii) civic friendship or (iv) complete harmony of persons, of their affections within and choices without.⁹² But they do aim at peace understood principally as security. It follows that the sovereign's authorization extends only to acts that secure and maintain peace understood in contrast to acts that destroy security. Therefore, by the terms of the sovereign-making covenant, a man or an assembly acts consistently with the sovereign power—acts *as a sovereign*—when he (it) acts for the sake of peace.

On this reading, the sovereign's power remains absolute. But absolute sovereignty does not entail the authority to command any member O2. Within the notion of absolute sovereignty is built the limitation of what gets to count as an *act of sovereignty*—and acts of sovereignty are always equitable, since equity is necessary constituent of the security of peace.

It should now be apparent how we can clarify Hobbes's account of absolute sovereignty to be compatible with the inalienable right to self-defense. When a sovereign orders one to perform acts destructive of one's life or judges iniquitously, the ordinance fails to be a binding command because the performance of such acts can never conduce to one's security. Such ordinances are *inconsistent with sovereignty*—call these IWS ordinances.

We can now assess the upshot of this account of sovereignty for Hobbes's understanding of civil law. If and when a sovereign dictates an

⁹¹ *Id.* at 17.1, 106.

⁹² For this partition of the degrees of the unity of peace, I borrow from Michael Pakaluk, *Is the Common Good Limited and Instrumental?* 55 REV. METAPHYSICS 57-94 (2001).

IWS ordinance, it fails to have moral validity because the subject has not actually authorized it. The reason the subject has not authorized it is because covenanters are taken to be practically reasonable in that they take good of life as basic and erect a sovereign for the sake of security. Commands to perform acts that would likely entail the destruction of one's life are just the sorts of ordinances that manifestly do not secure one's person, just as judgments that issue in inequity are bellicose. Since the sovereign is authorized only insofar as his acts secure one's person, and the common good of security, it follows that such ordinances are not members of the set O1—those ordinances that the sovereign is able to command *as sovereign*. In this light, consider the following passage:

The obligation of subjects to the sovereign, is understood to last as long, and no longer, than the power lasteth, by which he is able to protect them. For the right men have by nature to protect themselves, when none else can protect them, can by no covenant be relinquished. The sovereignty is the soul of the commonwealth; which once departed from the body, the members do no more receive their motion from it.⁹³

This passage can easily be read to indicate the notion of sovereignty as the mere fact that a state has sufficient guns, police, and military forces. But on the foregoing interpretation, those members of commonwealth who are addressees of a command in contravention to the basic precept of natural law to preserve one's life are addressees of commands inconsistent with sovereignty. Hence they are effectively not moved by the sovereign's command, as members of a body are not moved by the soul departed.

It is an interesting question as to the membership status in the commonwealth of the addressee of a morally invalid law. Is the addressee still a member of the commonwealth, retaining a right to *resist* the sovereign? Or is the addressee of an IWS ordinance thrown back into the state of nature, in which case one's former sovereign is now a very powerful enemy? Those who defend the latter contend that, in commanding what I have called an IWS ordinance, the sovereign thereby violates the covenant—and the very being of the commonwealth presupposes the covenant. So, commanding such an ordinance, the entity that was the commonwealth is no longer—at least for the addressee. At best it is an entity claiming to be a commonwealth. The strength of this solution is that it retains the absolutism of the commonwealth or state—*so long as it qualifies as a state*.⁹⁴ It would retain Hobbes's apparent commitment to forestalling all rebellion. Moreover, it takes seriously Hobbesian reckoning of the meaning of words.

Still, those who defend the former solution maintain that the right to self-defense is properly a right of resistance *against the state*. On this reading, addressees of IWS ordinances remain members of the commonwealth, but are justified in disobeying such commands. In favor of this solution is, chiefly, that Hobbes refers to self-defense rights as liberties *of subjects*. So

⁹³ L, 21.21, 144.

⁹⁴ See Peter J. Steinberger, *Hobbesian Resistance*, 46 AM. J. POL. SCI. 856-65 (2002).

someone commanded to incriminate himself remains a member of the commonwealth—but has a right to disobey it. This may be taken to imply that addressees of commands IWS remain subjects.⁹⁵

My argument is compatible with either solution. Under the first solution, an addressee of an IWS command rightly views it as not only morally invalid, but systemically invalid as well, since the addressee would thereby no longer have a rule of recognition. Non-addressees of the command, inasmuch as they rightly reason, will recognize the command as systemically valid but morally invalid. Under the second solution, both addressees and non-addressees of an IWS command view it as morally invalid, but recognize its systemic validity. If one remains a subject, one can recognize the systemic and moral validity of other laws of the sovereign.

What is the status of a morally invalid command, then? Is it a *law*? Given Hobbes's claim that it is of the very nature of law to bind—*pace* latter-day positivists, Hobbes denies that law merely “claims” to bind—then we must deny it the status of civil law for the addressee, because such commands do not bind one to act. Civil laws, properly speaking, must proceed from acts of the sovereign as sovereign—and morally invalid commands do not proceed from the commander as sovereign. Still, this may be compatible with the addressee recognizing that some such edict has a measure of the character of law, inasmuch as it is systemically valid, if we take the addressee to still be a member of the commonwealth.

To sum up the solution to the obscurity of systemic and moral validity apparent in Hobbes's theory of authorization: persons covenant to authorize a sovereign for the sake of security. By its very nature, the covenantal act authorizes only those acts consistent with sovereignty—and I have argued that Hobbes builds into the notion of sovereignty not only the rights essential to execute its end, but also the end for which the covenant was made. This is how Hobbes can have an absolute sovereignty that is compatible with inalienable rights to self-defense. Accordingly, this is how Hobbes can at once have an absolutist understanding of civil law while being able to retain content-based limitations on which ordinances can achieve the status of civil *law* (i.e., commands that of their nature bind in conscience). Indeed, this interpretation is supported by the so-called mutual containment thesis—the point of this thesis is to secure the practical congruence between civil law and natural law. Let us turn now to consider it in detail.

VI. THE MUTUAL CONTAINMENT THESIS AND THE SOVEREIGN RIGHT OF JUDGMENT

Recall that the second way Hobbes appears to obscure his properly natural law account of civil law is in his claim that “law of nature, and the

⁹⁵ SREEDHAR, *supra* note 70, at 156-57.

civil law, contain each other, and are of equal extent.”⁹⁶ We have already suggested that this means at least that the sovereign is erected in order to provide the security condition for *in foro externo* validity of the laws of nature. The laws of nature will only bind *in foro interno* until a sovereign is erected to sanction noncompliance; if they bound *in foro externo* prior to the sovereign-making covenant, they would make one prey to others, contrary to the more basic rational necessity to preserve one’s life. I shall argue that the mutual containment thesis actually supports my interpretation that there are content-based limitations on what can be effected into civil law, i.e., moral validity is a necessary condition for commands and judgments to bind subjects.

I alluded earlier to a possible interpretation that could present a challenge to my argument that Hobbes’s account of civil law is properly speaking a natural law account. I have in mind the recent work of Lloyd, who has argued that, while the laws of nature have a normative status prior to civil law, natural law is “self-effacing” and Hobbes is a practical legal positivist. This interpretation, Lloyd claims, is the key to understanding the mutual containment thesis. Notably, Lloyd’s positivist interpretation of Hobbes is different from the standard positivist interpretations. The standard interpretations tend to trade on one of the secularist theses regarding God’s nonessential role in Hobbes’s legal philosophy: the historical, concealment, and practical severability theses. Lloyd’s argument explicitly avers that her interpretation is compatible with theistic interpretations of the laws of nature as really laws.⁹⁷ I shall argue, on the basis of Hobbes’s theories of equity and right judgment, that Hobbes is not a practical legal positivist. If successful, my argument will show how the natural law reading of Hobbes is incompatible with a positivist reading of his theory of civil law.

On Lloyd’s reading the state of nature is a condition in which all persons have an absolute right of private interpretation about what the natural law requires. This condition is effectively a powder keg that all too quickly devolves into war. Since they desire peace, persons will want others to submit their right of private judgment to a political authority, and according to Hobbes’s “easy sum” of the laws of nature it follows they ought to as well. The “easy sum” is: *Do not that to another, which thou wouldst not have done to thyself*.⁹⁸ Lloyd reformulates the easy sum as the Reciprocity Theorem (RT): *to do what one condemns in another is contrary to reason*.⁹⁹ RT requires that persons in the state of nature give up their right of private judgment and set up a supreme judge and arbiter. Thus, for Lloyd’s Hobbes, judicial supremacy is the essential feature of sovereignty.

⁹⁶ For an interesting account of the mutual containment thesis, which seems compatible in many ways with my own account, see Timothy Fuller, *Compatibilities on the Idea of Law in Thomas Aquinas and Thomas Hobbes*, 112-34 HOBBS STUDIES III (1990).

⁹⁷ LLOYD, *supra* note 15, at 265.

⁹⁸ L, 15.35, 99.

⁹⁹ LLOYD, *supra* note 15, at 219-20.

The sovereign is empowered to “legitimately settle disputes as to *what the law—including natural law—is*.”¹⁰⁰ This is how Hobbesian natural law is “self-effacing”: natural law itself requires people to give up their right of interpreting the meaning of natural law to an all powerful judge. In this way, no one can pretend to disobey sovereign judgments on the basis that the sovereign’s judgment conflicts with natural law. This is Lloyd’s explanation of what it means for the civil law to contain the natural law.¹⁰¹

From these grounds, Lloyd infers that, “there is no legitimate position or perspective from which we can criticize or resist the sovereign’s decisions.”¹⁰² The sovereign’s judgment is authoritative even if or when it is “cosmically incorrect.”¹⁰³ Even if the sovereign’s judgment is erroneous, natural law directs us subordinate our judgment to him (it) or else risk the return to anarchy.¹⁰⁴ Whereas, my foregoing argument suggested that the natural law provides content-based criterion on what counts as acts of sovereignty, Lloyd claims that the natural law directs agents to completely subordinate private judgment to an unlimited sovereign judge—even when it is “cosmically” incorrect, which on Hobbes’s terms would mean nothing other than contravention of God’s will manifest in natural law.

I maintain that such an interpretation at once fails to correctly understand Hobbes’s theory of equity as a moral check on the sovereign and the real import of the mutual containment thesis. Let us consider Hobbes’s notion of equity more closely.

We saw earlier how Hobbes maintained, on his theory of authorization, that nothing the sovereign can do is “unjust.” Our reconstruction of Hobbes suggested that this should be understood of the sovereign insofar as he (it) acts as sovereign or according to the end of the sovereign office. In this way, a command contravening the basic rational necessity to preserve one’s life would not be acting *as sovereign*, and so technically, “justice” is not implicated because the addressee’s exercises his right to self-defense outside the confines of covenant. While Hobbes won’t admit the sovereign can do “injustice,” he does say that the sovereign “may commit iniquity.”¹⁰⁵

According to Hobbes, the “general rule of equity” is both “the law of reason” and “the law of God.”¹⁰⁶ The eleventh law of nature in *Leviathan’s* numbering, equity consists in dealing equally when judging between man and man or the “equal distribution to each man, of that which in rea-

¹⁰⁰ Lloyd, *supra* note 80, at 295.

¹⁰¹ For a discussion of Lloyd’s interpretation of the mutual containment thesis and Hobbes’s theory of equity in a different context (as it might apply to contemporary U.S. Supreme Court abortion jurisprudence), see my *The Prolife Leviathan: The Hobbesian Case Against Abortion* 86 (4) AM. CATH. PHIL. Q. 557 (2012).

¹⁰² Lloyd, *supra* note 80, at 295.

¹⁰³ LLOYD, *supra* note 15, at 342.

¹⁰⁴ *Id.*

¹⁰⁵ L, 18.6, 113.

¹⁰⁶ EW VI, 21-2.

son belongs to him,” and is properly the act of a judge or arbitrator.¹⁰⁷ Like the easy sum, the natural law of equity flows from fundamental human equality, since disputes between equals merit impartial judgment. Hence, the judge who “performs his trust” deals equitably between persons. The judge who deals inequitably is “partial in judgment” and is a practitioner of “acceptation of persons.”¹⁰⁸

Hobbes thus recognizes the possibility that the sovereign judge can fail to act according to the precept of equity. As Noel Malcolm puts it, Hobbes’s notion of equity “shows that morality remains an objective standard, by which the laws or actions of the sovereign can be judged.”¹⁰⁹ My contention is that the law of equity binds the sovereign in a way that entails that mere systemic validity of some judgment is not sufficient to bind subjects. All iniquitous judgments are, properly speaking, inconsistent with sovereignty—and this broadens the class of IWS acts. Accordingly, all iniquitous judgments are morally invalid on the ground of the independent criterion of equity. Since systemic and moral validity are necessary conditions for some command or judgment to bind in conscience, and since it is of the very nature of law to bind, it follows that mere systemic validity is not sufficient for an act to count as a *law*. Hobbes’s theory of equity cannot be squared with a legal positivist interpretation of his theory of civil law.¹¹⁰

Hobbes’s condemnation of the iniquitous judgments of his day is suggestive of their moral invalidity. One example is Hobbes’s condemnation of a judgment set down by Sir Edward Coke. Coke had contended that an innocent man accused of a felony who flees for fear of corrupt or partial judges and who is afterwards brought to trial and proved innocent, shall, notwithstanding his innocence forfeit his goods and property.¹¹¹ Hobbes rejects as iniquitous the presumption that the flier is guilty. Such a presumption deprives an innocent man his due in violation of equity, because reason requires that innocent persons not be deprived of their goods or property, and equity binds judges to equally distribute “to each man, of that which in reason belong to him.”¹¹² The natural law of equity brooks “no exception at all.”¹¹³ As long as such a precedent stands by the sovereign’s will or tacit permission, the law is systemically valid (and, in some cases that is enough to garner merely prudential obedience out of respect

¹⁰⁷ L, 15.24, 97.

¹⁰⁸ *Id.* at 15.24, 97.

¹⁰⁹ NOEL MALCOLM, ASPECTS OF HOBBS 437 (2002). While Malcolm helpfully illuminates the sovereign’s office to promote the good of the people, my argument suggests Malcolm is incorrect that the sovereign is “jurally entitled to treat them just as he would his enemies,” if we take “civil law” to be a jural term (446).

¹¹⁰ As, for example, Zagorin argues. See ZAGORIN, *supra* note 13, at 92-95.

¹¹¹ L, 26.24, 181-2; Cf. EW VI, 137-8 and EDWARD COKE, THE FIRST PART OF THE INSTITUTES OF THE LAWS OF ENGLAND, §709 (8th ed., 1670).

¹¹² L, 15.23, 97; EW VI, 21-2.

¹¹³ EW VI, 137.

for the law as a whole, including the equitable parts of the law). But, the upshot of Hobbes's criticism is that it is morally invalid.

One immediately notices that, in criticizing the common law of his day, Hobbes himself does not practice the sort of absolute subordination of judgment that Lloyd claims Hobbesian natural law demands. Neither can we say that Hobbes's criticism was of Coke and not of the sovereign, since Hobbes maintains that common law has the force of civil law by the sovereign's tacit will. But, let us set that point aside. The natural law of equity is supposed to be a principle binding the judge and that the judge can fail to judge equitably. Lloyd maintains that when judges fail, Hobbesian natural law unflinchingly binds us to obey, because resorting to the right of private judgment would risk return to anarchy.

But, consider Hobbes's remark in his discussion of the natural law of equity that the acts of the inequitable judge are the "cause of war."¹¹⁴ As Lloyd correctly points out, the sovereign is the supreme judge. It follows from Hobbes's remark that the sovereign will can fail to be equitable when it fails to perform its trust—and when he (it) fails to judge equitably, he (it) causes war. Because inequitable judgments are always bellicose, they are always inconsistent with sovereignty, since the very purpose of the sovereign office and power is the peace of security.

Why does the failure to judge equitably cause war? Hobbes answers that the inequitable judge "doth what in him lies, to deter men from the use of judges, and arbitrators."¹¹⁵ It is fairly obvious why the deterrence of men from the use of judges brings on a state of war for the same reasons Lloyd gives: erecting a sovereign judge entails giving up the right to be judge in one's own case. This means that one gives up the *right of private judgment* over good and evil. So, to deter men from the use of judges and arbitrators would be to encourage men to rely on the right of private judgment. But the widespread practice of the right of private judgment would inaugurate war. As a potential condition of actual persons, the state of nature is a *reductio* of the practice of the right of private judgment on a massive scale.

Yet, Lloyd's claim that the laws of nature direct subjects to completely subordinate their judgment to the sovereign judge on the pain of the subjects causing war (or initiating a chain of causes leading to war)—even when his (its) judgments are "cosmically" iniquitous—ignores Hobbes's point that it is *the sovereign* who causes war when he (it) judges iniquitously. The *sovereign* himself (itself) *causes* a state of war by *detering* agents from the use of judges. How is that?

The sovereign's iniquitous judgment need not initiate *actual* fighting or battle for it to cause war, according to Hobbes. War consists not only in actual fighting, but also "in the known disposition thereto, during all the

¹¹⁴ L, 15.23, 97.

¹¹⁵ *Id.*

time there is no assurance to the contrary.”¹¹⁶ Suppose there is a case in which the sovereign judges iniquitously and Smith discerns that he has been unreasonably harmed. Suppose further that Smith is a reasonable fellow and he acts according to RT: he does not do that which he condemns in another. (These suppositions are warranted—when Smith becomes a *subject*, he does not become a new kind of thing—he is still a *rational animal*.¹¹⁷ Hence, he has not lost his reasoning powers in erecting a sovereign judge.)¹¹⁸ Smith is now doubtful that he will get a fair shake from appealing to the sovereign judge. But Smith is not the only one. Insofar as iniquitous judgments of the sovereign are publicly known, others are deterred from the use of judges, too—and Smith knows that. Smith knows that his neighbor, Jones, is not assured of getting a fair shake from the sovereign judge. Hence, Smith knows Jones is deterred from appeal to the sovereign judge and encouraged to assume a right of private judgment—and he wouldn’t blame him if he did. After all, Smith was the victim of the iniquitous judgment and he wouldn’t blame his neighbor if his neighbor judged that his person, family, and goods were no longer safe from iniquity in the sovereign’s court. If we accept the RT as a valid formulation of the easy sum, then, if Smith is deterred from appeal to the sovereign judge (and accordingly assumes a right of private judgment), he acts reasonably, since he does not condemn that in Jones. This story shows, *pace* Lloyd, how the easy sum or its formulation in RT does not direct unfailing subordination of judgment to the sovereign.

Given the sovereign’s iniquity, and his neighbor’s knowledge of it, Jones is reasonable to judge that the security of his person, family, property, etc. are in jeopardy not only from the sovereign, but from his neighbors. Hobbes explicitly warns that corrupt judgments will lead to this chaotic state of affairs. While this statement comes in the context of explaining what sovereigns should cause to be taught to subjects, Hobbes is actually teaching about the duty of the sovereign in his capacity as judge, and may even be suggesting subjects should remain vigilant to watch out for corruption in the sovereign’s courts:

For which purpose also it is necessary they be showed the evil consequences of false judgment, by corruption either of judges or witnesses, whereby the distinction of propriety is taken away, and *justice becomes of no effect*: all which things are intimated in the sixth, seventh, eighth, and ninth commandments.¹¹⁹

It follows that if the sovereign judges iniquitously at t_1 then, the *sovereign* has initiated a time series (t_1, t_2, t_3, \dots) in which men lack the assurance of peace—iniquitous judgment inaugurates a condition of war until such time as peace is reassured.

¹¹⁶ *Id.* at 13.8, 76.

¹¹⁷ EW IV, 2; I, 136.

¹¹⁸ Fuller makes a similar point. See Fuller, *supra* note 96, at 114-15.

¹¹⁹ L, 30.12, 224 (emphasis added).

Since the sovereign's iniquitous judgment causes war, the absolute subordination of private judgment to the sovereign judge does not seem to be the meaning of the mutual containment thesis. I suggest that the true import lies in Hobbes's indication that commands and judgments fail to have moral validity because they incite war, and that this is illuminated in Hobbes's *in foro interno-in foro externo* distinction.

We have already seen how the *in foro interno-in foro externo* distinction is supposed to make sense of how Hobbes's catalogue of natural laws can be immutably and eternally binding in a way that does not require self-destruction. The catalogue of natural laws do not always bind *in foro externo* in the state of nature. But this does not entail that they lack legal force by their divine pedigree. It only means that the laws of nature do not require one to unduly risk getting double-crossed. The situations Hobbes has in mind where performance of a contract would lead to one's own destruction are those contracts and covenants formed in a condition in which there is not a common power to provide sanctions for breach of the natural law. This is the condition of war.

When the sovereign judges iniquitously, he (it) causes war. Hobbes seems to mean that, when the sovereign commands or judges contrary to basic equity, the sovereign has initiated a state of affairs in which men and women would be justified in judging that there is no longer "sufficient security" to observe the laws of nature *in foro externo*. We have already seen why. When the sovereign publicly judges iniquitously, he deters reasonable persons from appeal to him (it) as judge. This has the potency to initiate a chain of causes in which reasonable persons are deterred from adverting to the common judge and are encouraged to assert a right of private judgment as a surer means to secure their persons, families, and goods. In this situation there is no longer mutual assurance of *in foro externo* obedience to the laws of nature, which just is the state of war. The state of war, as distinct from a state of peace, may be either a *disposition toward* or *actual* assertion of the right of private judgment on a massive scale. Either way, such a condition is not a peaceful one. The non-peaceful condition is one in which, by definition, sufficient security fails to obtain for the laws of nature to bind *in foro externo*. But the *in foro externo* "putting into place" of the laws of nature just does consist in obeying (at least a major part of) the civil law, since it contains the natural law. Therefore, the warlike condition is one in which the civil laws do not bind one *in foro externo*.

It follows that the sovereign's own iniquitous judgment fails to attain the status of binding civil law because in the very act of judging in this way, the sovereign brings on a state in which there will be a tendency to the widespread assertion of private judgment and fisticuffs. In that case, one would not be assured of the *in foro externo* compliance of others—and the only precept that would be binding both *in foro interno* and *in foro externo* is the basic rational necessity to preserve one's life. This is another way of saying that for something to become civil law it must be both systematically and morally valid. It must be congruent with natural law which, as in the classical natural law tradition, is ultimately congruence with God's will.

VII. CONCLUSION

I have argued that Hobbes should be interpreted to give a properly natural law account of civil law. This means that assessing the lawful character of some ordinance cannot rest simply on its human source or pedigree, independent of (or separable from) its moral content. For Hobbes, the moral criterion is rooted in his natural law theory of morality, the heart of which is the basic good of life and security. For something to bind with the force of civil law it must give the addressee a sufficient reason to act—it must be both systemically and morally valid. Hobbes's theory of authorization and the mutual containment thesis have been taken to collapse systemic and moral validity under the sovereign will. This is in a sense true, but only inasmuch as the sovereign acts as a sovereign—and that “inasmuch as” imports the ends of the sovereign-making covenant—life, security, and equity—as validating conditions of which possible ordinances can count as civil law.

When seen in this light, we can start to see more clearly how Hobbes could maintain that civil law and natural law are necessarily congruent just as he can maintain a normatively charged notion of sovereignty-for-the-sake-of-security. Natural law requires us to secure our persons. Hence, it requires us to erect a sovereign powerful enough to protect us. If and when the sovereign acts violently toward those who authorized him (it) to secure the peace, the sovereign fails to bind addressees with the force of civil law. Hobbes has effectively ensured the practical congruence of natural law and civil law because for an ordinance to attain the full status of civil law, it must be systemically and morally valid. This account is consistent with Hobbes's absolutist conception of the sovereign's lawmaking and adjudicative powers because morally invalid commands and judgments fail to be consistent with sovereignty.

The foregoing is of course only an outline of a natural law interpretation of Hobbes's theories of morality and civil law. But, if the account has merit, then Hobbes's place in standard narratives of the history of natural law theory and jurisprudence needs to be reassessed. Indeed, Hobbes stands at the fulcrum between premodern and modern natural law theory. But, it may be that Hobbes's break with classical natural law theory lies neither in its supposedly secular foundations, nor in a proto-Humean account of practical reason, nor in a positivist account of civil law. Rather, Hobbes's break from the older tradition may consist chiefly in lopping off those goods that correspond to the higher levels of Aquinas's metaphysical stratification, *qua* objective. If correct, then Hobbes's distinctive contribution to natural law theory seems to lie in his thin theory of the good for assessing the validity of civil law.

MELTING POT BENEVOLENCE AND LIBERTY
PATRIOTISM: THE IMPORTANCE OF THE MORAL
COSMOPOLITANISM PRECEDENT IN ASIAN AMERICAN
HISTORY

Will Sarvis*

ABSTRACT

Between the 1860s-1930s, there were a significant number of Chinese, Japanese, and various Caucasian peoples who embraced interracial friendships in the United States. Not only were these brave souls ahead of their time, but they exercised a moral cosmopolitan attitude amidst some of the fiercest racial discrimination in American history. Until recently, racism, exclusion, and ethnic discrimination have understandably dominated the historiography of Asian America. Scholars have also thoroughly documented formal legal remedies that proponents of interracial benevolence sought, largely in federal courts. The moral cosmopolitan precedent, on the other hand, offered an antidote to racism largely outside the confines of systematic law. In the tradition of populist constitutionalism, the moral cosmopolitan precedent of pre-Depression interracial goodwill among Asians and Caucasians created a minority societal dissent that later became the broad cultural basis for overturning institutionalized discrimination. This essay makes the case for legal and extra-legal moral cosmopolitanism, where a Melting Pot version of multiculturalism merged with a patriotism that embraced Enlightenment principles of universal human rights.

CONTENTS

I.	INTRODUCTION	198
	A. NOMENCLATURE.....	199
	B. HISTORICAL MORAL COSMOPOLITANISM.....	199
	C. THE DREAM AND REALITY OF A NATION OF IMMIGRANTS.....	200
II.	BEFORE THE PROGRESSIVE ERA.....	202
III.	UTILITARIAN FRIENDSHIPS DURING THE PROGRESS- IVE ERA AND BEYOND.....	209
IV.	PROGRESSIVE ERA FRIENDSHIPS OF PLEASURE AND VIRTUE	224
V.	THE IMPORTANCE OF THE MORAL COSMOPOLITAN PRECEDENT	240

VI. LIBERTY PATRIOTISM IMPLICATIONS254
VII. CONCLUSION256

I. INTRODUCTION

This study focuses on interracial goodwill during a time and place of intense racism and nativism, that of the United States during the 1850s-1920s. Part of its historiographical significance lies in its unprecedented telling. Asian-American historiography has focused mainly on racism, discrimination, and exclusion.¹ In 2001, Gary Okihiro wrote that “Anti-Asianism is the most discussed topic in Asian American history,”² with the Japanese-American relocation during World War II probably being the most-discussed topic within Anti-Asianism.³ There is good reason for this. The Chinese were the first immigrants to be excluded by race from the United States. Japanese-Americans suffered uniquely during World War II with forced relocation to internment camps, with resultant loss of income, property, and social relationships, not to mention peace of mind. Asian immigrants were the only ethnicity that the U.S. government completely excluded from immigration or banned from naturalization.⁴ Scholars have very thoroughly documented these sorts of stories. In recent years the field of Asian American history has finally begun to move in new directions, but the old beating boy of the evil trio (racism, discrimination, and exclusion) lives on. This work aims to contribute to the corpus of work that is now developing beyond the old evil trio.

A. NOMENCLATURE

In this essay I will use terms like “Asian” and Caucasian” in somewhat inexact ways, but compatible with racial views of the nineteenth and early twentieth centuries based almost exclusively upon physical appear-

*Will Sarvis is a freelance writer. The author kindly thanks Anne Richardson Oakes, an anonymous referee, and a contingent of law students for helping bring this work to publication. This essay is dedicated to three of the author’s former professors, moral cosmopolitans all: David Burr, Larry Shumsky, and Young-tsu Wong.

¹ Early scholars working in Asian American history were naturally preoccupied with documenting racism, exclusion, and discrimination against people of Asian descent. See HYUNG-CHAN KIM, ASIAN AMERICAN STUDIES: AN ANNOTATED BIBLIOGRAPHY AND RESEARCH GUIDE Sections 6, 7, 13 (1989). Unfortunately, this remains the preoccupation of the field. Conflicts, past and present, predominate in JEAN YU-WEN SHEN WU & THOMAS C. CHEN, ASIAN AMERICAN STUDIES NOW: A CRITICAL READER (2010).

² GARY Y. OKIHIRO, THE COLUMBIA GUIDE TO ASIAN AMERICAN HISTORY 75 (2001).

³ *Id.* at 100.

⁴ FRANKLIN ODO (ed.), COLUMBIA DOCUMENTARY HISTORY OF THE ASIAN AMERICAN EXPERIENCE 1 (2004).

ance. Therefore I will not explore contemporary ideas which rightfully dismiss any scientific basis for “race,” and see it for the social construct that it is and was. Unfortunately, too few hold these post-racial ideas now, but they would have been especially irrelevant to the vast majority of people living in the United States a century and more ago. Also, for convenience sake and because of dramatic demographic differences, this study concentrates on the U.S. mainland, most notably excluding Hawaii. Hawaii did not become an American territory until late in this study’s chronology (1898), but also featured proportionately far more Asian and indigenous peoples than the mainland. Race relations there deserve more particular, more focused studies that would likely contrast in major ways from the one offered here.

B. HISTORICAL MORAL COSMOPOLITANISM

Long before scholars re-discovered the philosophy,⁵ these interracial friends embraced an *attitude* of moral cosmopolitanism.⁶ They were not necessarily intellectually sophisticated people, nor were they likely to be self-consciously “citizens of the world” in an educated understanding of cosmopolitanism.⁷ But obviously race had not prevented the development of intimate friendships among them, and so they must have appreciated people unlike themselves, at least to some extent, as part of a greater family of humanity.⁸ In describing true friendships, Aristotle wrote “loving resembles an affection of the soul, whereas friendship resembles a disposition . . . reciprocal loving involves decision, and decisions flow from dispositions, and when people wish good things for those they love for these others’ sake, this is not a matter of affective state but of disposition.”⁹

⁵ Early references appear in John HIGHAM’s classic study of nativism. See JOHN HIGHAM, *STRANGERS IN THE LAND: PATTERNS OF AMERICAN NATIVISM 1860-1925* 22, 122, 125 (1975). For more recent works, see Jeremy Waldron, *Minority Cultures and the Cosmopolitan Alternative*, 25 U. MICH. J.L. REFORM 751 (1992); DAVID A. HOLLINGER, *POSTETHNIC AMERICA: BEYOND MULTICULTURALISM* 3-5 (1995); Robert Fine & Robin Cohen, *Four Cosmopolitan Moments*, in *CONCEIVING COSMOPOLITANISM: THEORY, CONTEXT AND PRACTICE* 140 (Steven Vertovec & Robin Cohen, eds., 2003); RICHARD THOMPSON FORD, *RACIAL CULTURE: A CRITIQUE* 163-66 (2005); see generally KWAME ANTHONY APPIAH, *COSMOPOLITANISM: ETHICS IN A WORLD OF STRANGERS* (2006).

⁶ APPIAH, *supra* note 5, at xv; Steven Vertovec & Robin Cohen, *Introduction: Conceiving Cosmopolitanism*, in *CONCEIVING COSMOPOLITANISM: THEORY, CONTEXT AND PRACTICE* 13 (Steven Vertovec & Robin Cohen, eds., 2003).

⁷ FORD, *RACIAL CULTURE*, *supra* note 5, at 164; APPIAH, *supra* note 5, at xiii.

⁸ An excellent article on the history of cosmopolitan philosophy, from its Greek roots to the present time, may be found in the Stanford Encyclopedia of Philosophy. See Pauline Kleingeld & Eric Brown, *Cosmopolitanism*, *STANFORD ENCYCLOPEDIA OF PHILOSOPHY* (Jul. 1, 2013), <http://plato.stanford.edu/entries/cosmopolitanism>.

⁹ NICOMACHEAN ETHICS (hereafter NE) § VIII.5, (Sarah Broadie & Christopher Rowe trans., Oxford Univ. Press 2002).

Disposition, inclination, or attitude thus becomes a core factor in a moral cosmopolitan worldview.

For purely organizational purposes, we might adopt Aristotle's ideas of tri-part friendships involving utility, pleasure, and deep, true friendship.¹⁰ Probably outside of Aristotle's tri-part categorization of friendships were the ambiguous relations fostered by churches, missionaries, charitable efforts, and various Caucasians who were sometimes only nominally anti-racist on culturally-specific and subjective moral grounds. This becomes especially true when considering the nineteenth century and the early twentieth century, when pseudo-scientific ideas of race hierarchy and Social Darwinism encouraged the development of an unfortunate cultural imperialism.¹¹ Nevertheless, it might be worth noting those who at least *meant* well, even if their own self-conceptions often colored their efforts. Cultural biases aside, pastors and churches were prominent in opposing enactment of the notorious 1882 Chinese Exclusion Act,¹² legislation that is easy to condemn now, but was unpopular to oppose then.

C. THE DREAM AND REALITY OF A NATION OF IMMIGRANTS

The United States is overwhelmingly a land of immigrants. The positive mythology surrounding this fact is one of multi-ethnicity and inclusiveness, the "bring me your huddled masses" inscription on the Statue of Liberty in New York City, supposedly protected by the Enlightenment era principle of all men (and eventually all women) being created equal. However, the dark reality of American history includes racism and nativism. Euro-Americans especially made aborigines, Africans, and their descendants the main targets of their racism, stealing land, resources, and labor and liberty from these peoples. During the late nineteenth and early twentieth centuries, nativists broadened the spectrum of discrimination to include anyone except themselves: white, Anglo-Saxon Protestants, the WASPs of American legend or notoriety.¹³ But regardless of specifics, racism has been a constant feature of American culture. Stricter immigration laws leading up to the 1924 Johnson-Reed Act¹⁴ mitigated nativist bigotry, but only because it assuaged their fears of being overwhelmed by foreign transplants. Eventually, the relegation of WASP ideology to numerical minority status,

¹⁰ *Id.* § VIII.3 .

¹¹ JOHN HIGHAM, STRANGERS IN THE LAND: PATTERNS OF AMERICAN NATIVISM 1860-1925 273 (1975); HYUNG-CHAN KIM, A LEGAL HISTORY OF ASIAN AMERICANS, 1790-1990 112 (1994).

¹² See the petition introduced by Massachusetts Senator George F. Hoar on behalf of various New York churches, 47 Cong. Rec. 2229 (1882); and the petition introduced by Ohio's Representative J. Warren Keifer on behalf of New York's east conference of the Methodist Episcopal Church, 47 Cong. Rec. 2840 (1882).

¹³ RICHARD HOFSTADTER, SOCIAL DARWINISM IN AMERICAN THOUGHT 172, 174 (George Braziller, Inc. 1959) (1944).

¹⁴ Pub. L. 68-139, 43 Stat. 153 (May 26, 1924). Also called the National Origins Act or the Immigration Act of 1924.

particularly after the influx of millions of Catholics, rendered nativism (ironically) irrelevant.

The truism of the United States being a land of immigrants is better known than its correlative, which is how successive generations tried to slam the door on further immigration. John Higham, in his classic study of nativism in general, *Strangers in the Land*, also described the anti-Chinese agitation of the 1870s and 1880s as the most violent expression of a long and varied attitude of American bigotry.¹⁵ The 1882 Chinese Exclusion Act became the only anti-immigration act in U.S. history to target a specific racial or ethnic group.¹⁶ Violence continued against the Chinese during the 1880s in a number of riots.¹⁷ California's Alien Land Laws of 1913¹⁸ and 1920¹⁹ targeted Japanese and attempted to prevent them from owning land. The Japanese ultimately experienced one of the most flagrantly racist violations of constitutional principles with their wartime internment.²⁰ Still, despite such prevalent racism, exclusion, and discrimination, there were numerous examples of interracial benevolence. In fact, the era's dominant racism and nativism makes such friendship all the more important when considering the aspirations of a multi-ethnic society such as the United States was or at least has come to be. Following is but a small sampling of the innumerable interracial friendships that existed among Asians and Caucasians before the Great Depression.

¹⁵ HIGHAM, *supra* note 11 at 25. See also Eiichiro Azuma, *A History of Oregon's Issei, 1880-1952*, 94:4 OREGON HISTORICAL QUARTERLY 315, 335-36 (1993-1994); Malcolm Clark, Jr., *The Bigot Disclosed: 90 Years of Nativism*, 75:2 OREGON HISTORICAL QUARTERLY 109 (June 1974); ROGER DANIELS, *THE POLITICS OF PREJUDICE: THE ANTI-JAPANESE MOVEMENT IN CALIFORNIA AND THE STRUGGLE FOR JAPANESE EXCLUSION* (1966); Daniel P. Johnson, *Anti-Japanese Legislation in Oregon, 1917-1923*, 97:2 OREGON HISTORICAL QUARTERLY 176 (Summer 1996); LAWRENCE M. FRIEDMAN, *AMERICAN LAW IN THE 20TH CENTURY* 145 (2002).

¹⁶ *Chinese Exclusion Act*, 22 Stat. 58, c. 126 (May 6, 1882); entitled "An act to execute certain treaty stipulations relating to Chinese."

¹⁷ CHARLES J. MCCLAIN, *IN SEARCH OF EQUALITY: THE CHINESE STRUGGLE AGAINST DISCRIMINATION IN NINETEENTH-CENTURY AMERICA* 173-74 (1994).

¹⁸ 1913 Cal. Stat. 206. See also FRANK F. CHUMAN, *THE BAMBOO PEOPLE: THE LAW AND JAPANESE-AMERICANS* 46-51 (1976); ROGER DANIELS, *ASIAN AMERICA: CHINESE AND JAPANESE IN THE UNITED STATES SINCE 1850* 145 (1988); ROGER DANIELS, *THE POLITICS OF PREJUDICE: THE ANTI-JAPANESE MOVEMENT IN CALIFORNIA AND THE STRUGGLE FOR JAPANESE EXCLUSION* 58-64 (1966); KIM, *supra* note 11, at 126.

¹⁹ 1921 Cal. stats, lxxxiii. See also CHUMAN, *supra* note 18, at 76-80; DANIELS, *POLITICS OF PREJUDICE*, *supra* note 18, at 88-92; KIM, *supra* note 11, at 126.

²⁰ Commonly known as Executive Order No. 9066. See Exec. Order No. 9066, 7 Fed. Reg. 1407 (Feb. 25, 1942) available at <http://www.ourdocuments.gov/doc.php?flash=false&doc=74&page=transcript>

II. BEFORE THE PROGRESSIVE ERA

In 1882 Congress passed and the president signed the Chinese Exclusion Act,²¹ the only immigration act in United States history to target a specific ethnic or racial group. The act followed a period of labor agitation against the Chinese, most spectacularly perpetuated in California by Dennis Kearney (1847-1907) and his Workingman's Party.²² But labor competition with (mainly European) immigrant groups was a prominent theme in American nativism going back to the 1850s, and especially salient during the Progressive Era of the 1890s-1920s. For the Chinese to be excluded had to illustrate a more pure form of racial hatred. This makes friendships among Chinese and Caucasians all the more remarkable for the pre-1882 period.

The first huge wave of Asian immigration to the United States occurred in conjunction with the California gold rush of 1848-1850. Thousands of Chinese arrived (mainly from the impoverished and war-torn Guangdong Province) hoping to make their fortune and, for the most part, return to China. Few people, including the Chinese, got rich during the gold rush, and afterward Chinese began working in a number of professions. They worked in agriculture, as domestic servants, in Pacific coast fisheries,²³ and perhaps most notably, as railroad construction laborers building the most difficult and most dangerous leg of the America's first transcontinental railroad from Sacramento to Promontory Point, Utah.²⁴

In addition to whites who were personal and public friends of the Chinese, there were a few prominent Caucasians who fought anti-Chinese prejudice in the legal-political world. As early as 1868, Oregon's governor George L. Woods vetoed legislation that targeted Chinese for extra taxation,²⁵ a technique common in mining areas. In Idaho two federal officials — U.S. Marshal Henry W. Moulton and U.S. Attorney Joseph W. Houston — also opposed such as discriminatory taxation of the Chinese.²⁶ So even before 1882 there was a small minority of prominent white Americans who might be characterized, without too much exaggeration, as the forerunners of multi-culturists. There were many more who did not cut such a conspicuous public profile. But if these moral cosmopolitans, prominent and obscure alike, had a spokesman, it was the popular writer Mark Twain.

Mark Twain famously defended the Nevada Chinese both in newsprint and his book *Roughing It* (1872). Twain, one of the most underrated anti-racists in American history, was also a defender of Native Americans

²¹ Chinese Exclusion Act (May 6, 1882) 22 Stat. 58, c. 126.

²² MCCLAIN, IN SEARCH OF EQUALITY, *supra* note 17, at 79-80.

²³ SUCHENG CHAN, ASIAN AMERICANS: AN INTERPRETIVE HISTORY 33 (1991); SHIH-SHAN HENRY TSAI, THE CHINESE EXPERIENCE IN AMERICA 10-13, 19-24 (1986).

²⁴ TSAI, *supra* note 23, at 15, 17.

²⁵ LIPING ZHU, "A CHINAMAN'S CHANCE" ON THE ROCKY MOUNTAIN MINING FRONTIER, IN CHINESE ON THE AMERICAN FRONTIER 239 (2001).

²⁶ *Id.* at 240.

and an opponent of foreign imperialism. Twain championed the Chinese as “quiet, peaceable, tractable, free from drunkenness, and . . . industrious as the day is long.”²⁷ Further, Twain wrote, “A disorderly Chinaman is rare, and a lazy one does not exist.” Twain blamed anti-Asian agitation on lower class, ignorant whites²⁸ — an observation many other Caucasians echoed during the following decades. Periodically, over the last couple of decades, reactionary forces have denounced Twain as a racist for the language he used in *Huck Finn*. But surely this is a most superficial reading of Twain. Besides his other, precocious anti-racist writings,²⁹ Twain’s hero *Huck Finn* famously denounces his society and risks condemning himself when he decides to aid the slave Jim in his escape.³⁰ In any case, Twain might be considered a “policy” friend to the Chinese and other ethnic and racial groups, since he used his influential pulpit as a writer to propound his views. On a more immediate level regarding the Nevada Chinese of this era, Virginia City policeman George Downey gained a reputation for ignoring race in criminal matters. Downey made an almost novel attempt in working class circles to understand Chinese culture. Wealthy Chinese merchants pooled their money and bought him an expensive diamond ring as just one symbol of their friendship.³¹

In protesting the exclusion of Chinese court testimony in California based on ethnicity, the Reverend William Speer wrote in 1870 that for them to be “excluded on ethnological grounds, is simply contemptible on the eyes of men of science.”³² Speer continued, “if the Chinese are Indians, then we are Indians; if the Chinese are negroes [sic], then we are negroes [sic].”³³ Speer was a cosmopolitan. Like Mark Twain and a handful of oth-

²⁷ MARK TWAIN, *ROUGHING IT* 391 (Elizabeth Frank ed., Signet Classics 2008) (1872).

²⁸ *Id.* at 392. See also Mark Twain, *Disgraceful Persecution of a Boy*, in *THE COMPLETE ESSAYS OF MARK TWAIN* 7-10 (Charles Neider, ed., 1963). The latter was originally published in *The Galaxy* (May 1870).

²⁹ For one of the most prominent examples, see Twain’s *Pudd’nhead Wilson*, a detective story that also ridiculed the idea of 100% whiteness as the standard against which all other people were measured. MARK TWAIN, *PUDD’NHEAD WILSON* (Dover Publications 1999) (1894).

³⁰ Jabari Asim accurately assesses all this in JABARI ASIM, *THE N WORD: WHO CAN SAY IT, WHO SHOULDN’T, AND WHY* 107 (2007).

³¹ Russell M. Magnaghi, *Virginia City’s Chinese Community, 1860-1880 in CHINESE ON THE AMERICAN FRONTIER* 133-134 (Arif Dirlik ed., 2003).

³² WILLIAM SPEER, *THE OLDEST AND NEWEST EMPIRE: CHINA AND THE UNITED STATES* 627 (1870).

³³ *Id.*

ers pioneering the conception of cultural relativism,³⁴ Speer was in this regard ahead of his time.³⁵

Twain, Downey, and Speer were unusual for their era, particularly in regard to the majority of Chinese, who at this time were laborers. During the flush economy of the immediate post-Civil War years, white laborers were generally not averse to allowing the Chinese to perform the Herculean task of blasting a rail path through the Sierra Nevada Mountains. Economic downturns, such as the Panic of 1873, frequently transformed grudging tolerance into overt racial targeting.³⁶ Labor competition was an early and frequent source of interracial and interethnic strife, but the Chinese and later Asians always bore the potential burden of a different physical appearance, which evoked racism rooted in the usual fear of the unknown.

Successful business relations required and require much social interaction, and congeniality must be present if such relations are to prosper. An enormous amount of Caucasian and Asian testimony supports the historical existence of many utilitarian friendships of this sort. Some developed further into true friendships, but all displayed the basic good will required of Aristotle's utilitarian friendships. Business interests became some of the most prominent and numerous protestors against the 1882 Chinese Exclusion Act.³⁷

In 1876 and 1886, during a nadir of anti-Chinese activity in the West, a number of prominent California businessmen, attorneys, pastors, and various others gave congressional testimony defending and praising the Chinese.³⁸ Joseph A. Coolidge (Secretary of the San Francisco Merchant's Exchange), Solomon Heydenfeldt (former associate justice of the California Supreme Court), and many others offered lengthy pro-Chinese testimony.³⁹ Frederick W. Macondray, of Macondray & Company and a twenty-four year resident of San Francisco, testified, "From all our dealings with them here and in China I do not know any class of merchants, I think, who are more honest and upright or who have a better reputation for integrity than the Chinese."⁴⁰ Macondray, who had lost money in his business dealings

³⁴ Franz Boas (1858-1942) deserves much of the credit for developing such ideas in the field of anthropology. See MAUREEN A. FLANAGAN, *AMERICA REFORMED: PROGRESSIVES AND PROGRESSIVISMS, 1890S-1920S* 263 (2006); Stephen Jay Gould, *The Mismeasure of Man* 140 (Norton, 1996) (1981); HIGHAM, *supra* note 11, at 125, 153, 304.

³⁵ Speer, however, had the Protestant missionary's bias that blinded him from valuing Chinese religious and philosophical traditions. See SPEER, *supra* note 32 at 605-24.

³⁶ HIGHAM, *supra* note 11, at 18.

³⁷ See the petition representing many business interests submitted by New York Senator Elbridge G. Lapham, 47 Cong. Rec. 2878 (1882); and the petition submitted by Massachusetts Senator Henry L. Dawes, 47 Cong. Rec. 3076 1st Sess. (1882).

³⁸ *Memorial of the Six Chinese Companies: An Address to the Senate and House of Representatives of the United States* 44th Cong. (1877); *The Other Side of the Chinese Question: Testimony of California's Leading Citizens* 49th Cong. (1886).

³⁹ *Id.*

⁴⁰ *The Other Side of the Chinese Question*, *supra* note 37, at 23.

with whites, testified he had never lost money with the Chinese. Furthermore, Macondray did all his business with the Chinese through verbal agreements. There was no need for written contracts.⁴¹ The bygone, evocative saying, “Their word was their bond,” evidently fit the Chinese merchants. These testimonies were remarkable. Out of jealousy, prominent businessmen of this and following eras of nativism tended to subvert successful immigrant businessmen, especially Jews.⁴² But in this singular context, a prominent class of whites defended a class of immigrant merchants. Equally remarkable defense of the Chinese arose in legal circles.

Benjamin Brooks and Frederick Bee were attorneys who came to champion Asian rights beyond the limits of professional duty. Bee defended Chinese miners in Nevada as early as 1855.⁴³ In 1876 he became the official spokesman for the Chinese Six Companies during Congressional hearings that received a majority of testimony against Chinese immigration.⁴⁴ Later Ch'en Lan-pin, Chinese minister to the United States, appointed Bee as consul out of the San Francisco office.⁴⁵ Both Brooks and Bee recognized the important economic role Asians had played in western United States development, denounced illegal anti-Asian activity, but also came to admire their Asian clients.⁴⁶ Writing as early as 1877, Brooks said,

The charges against this people . . . would in other lands be esteemed virtues; their undying love for native land, their devotion to their religious faith, their veneration for their parents and ancestors, their love and affection for their families, their generous contributions to their support and happiness, their untiring industry, their uncomplaining patience, their courageous venturing to the most distant lands where honest wages may be earned, their frugality, proud independence, resistance to oppression and partial laws If they cannot assimilate with us, can it be because these qualities are foreign to *our* nature? [original emphasis]⁴⁷

Obviously Brooks had gained some appreciation for the virtues of Chinese culture, and spoke with an amazing insight that has yet to characterize an appreciable number of white Americans, these 137 years later.

Judges, whose non-profit role allowed for more potential purity in their motives, became among the most stellar legal allies for Asians. J.S. Look, a Seattle-based Chinese businessman, remembered the strong anti-Chinese feeling of 1880s. “[But] there were a few big men in Seattle who sympathized with the Chinese race and who did much to bring about an era of good feeling,” Look recalled in 1924. “Some of these men I remember by name were Judge Jacobs, Judge [Thomas] Burke, Judge [C.H.] Han-

⁴¹ *Id.*

⁴² HIGHAM, *supra* note 11, at 26-27, 66-67, 92-94, 160-161.

⁴³ McCLAIN, IN SEARCH OF EQUALITY, *supra* note 17, at 64.

⁴⁴ *Id.* at 86.

⁴⁵ *Id.*

⁴⁶ *Id.* at 64-65; 86.

⁴⁷ *Id.* at 310, n. 98.

ford, Mr. Denny and Mr. [H.L.] Yesler”⁴⁸ These and other prominent community leaders (such as Judge Roger S. Green, Justice of the Peace George G. Lyon, Reverend L.A. Banks, Seattle Fire Chief Gardner Kellogg)⁴⁹ with power and stature, had taken an unpopular stand during one of the most anti-Asian racist decades in West Coast U.S. history.⁵⁰

Those bent upon emphasizing ethnic and racial conflict might be tempted to say these white men were merely “law and order” types interested in preventing anarchy. But let us remember that a great many white authorities looked the other way (or covertly or overtly participated) in lawlessness directed against blacks in the U.S. South of the entire Jim Crow era (1877-1954) and beyond. In fact, when one racial group systematically allows domestic terrorism against another, such practices *become* a form of (racist) orderliness. So the Seattle authorities’ actions cannot be separated from the Chinese they sought to protect. They must be described as friends, albeit in a professional capacity executing their professional duty. Captain George Kinnear, leader of the Home Guards that helped quell the 1886 Seattle riot, recalled in 1911:

The deplorable situation and the cause of all our trouble was — two few men were willing to throw themselves into the breach to defend the right at any cost, and too many were afraid to do anything to check the tide of lawlessness. Professional men were afraid they would lose some of their clients. Merchants were fearful they would be boycotted. The merchants in the building in which Judge [Thomas] Burke had his office said he must vacate and leave the premises for fear the building would be fired [burned down] or blown up. But the Judge stayed. He was one of the men who put down the mob.⁵¹

Unsurprisingly, Judge Roger S. Greene’s wife started a Chinese school during this period to help teenagers like L.O. Dong (president of Seattle’s Chinese Association during the 1920s).⁵² These Caucasians cared about their community and its members, Asians and whites alike.

In Portland, Matthew P. Deady (Oregon’s first federal district judge) evolved from a pro-slavery advocate to one of the most vocal pro-Chinese

⁴⁸ J.S. Look Interview, SURVEY OF RACE RELATIONS, box 27, file 182, at 2 (on file with the Hoover Institution Archives, Stanford Univ.). Look is no doubt referring to some of those who, among other things, faced down the anti-Chinese rioters in Seattle in 1886. Upon the 25th anniversary of this sad event, George Kinnear (eye witness and law enforcement participant) published his memoir of the event. See GEORGE KINNEAR, ANTI-CHINESE RIOTS AT SEATTLE, WASHINGTON, FEBRUARY 8TH 1886 (1911).

⁴⁹ KINNEAR, *supra* note 48, at 6, 7, 10. KINNEAR names many others who risked their lives and enforced the law. For members of the 14th Infantry and Home Guard see KINNEAR, *supra* note 55, at 10, 16-17.

⁵⁰ William Henry White (1842-1914), U.S. District Attorney 1885-1889 under President Grover Cleveland, also appears to have joined the others in quelling the rioters. See SEATTLE DAILY TIMES (Apr. 29, 1914) at p. 3, c. 1.

⁵¹ KINNEAR, *supra* note 48, at 13-14.

⁵² Long O. Dong Interview, SURVEY OF RACE RELATIONS, box 27, file 171, at 2.

advocates.⁵³ Perhaps most outstanding was Judge Ogden Hoffman, who heard over 10,000 habeas corpus cases alone following the 1882 Exclusion Law.⁵⁴ This was not surprising to anyone who had observed Hoffman's earlier judicial performance. Between 1854 and 1872, California law forbade the admission of Chinese testimony in state courts. In Hoffman's federal court, however, Chinese were allowed to bring suit against fellow Chinese and non-Chinese alike. During Hoffman's entire tenure only two such plaintiffs received unfavorable judgments. Chinese convicted of crimes also generally received comparatively lenient sentences in Hoffman's court.⁵⁵ Clearly Hoffman and other federal judges of the northern California district were at least legal friends to the Chinese, despite holding personal racial biases that were common at the time. Apparently here, the institution of the law itself, particularly regarding habeas corpus and due process, became the driving forces behind the judges' decisions.⁵⁶ Hoffman maintained an alternative to California's shameful testimony prohibition as sort of a running dissent, even if in his limited judicial arena. Eventually the state legislature amended the California's civil and penal codes to reinstate Chinese testimony.⁵⁷

The prohibition against Chinese testimony in California state courts had its roots in 1849-50, when the state legislature passed an act that barred non-whites from giving evidence for or against whites. But the legislature only specified Indians and those of African ancestry.⁵⁸ The specification against the Chinese arrived a few years later with the infamous 1854 California Supreme Court case, *People v. Hall*.⁵⁹ A lower court had convicted one George W. Hall for murdering a Chinese man named Ling Sing and sentenced him to execution. Hall appealed upon the openly racist grounds that testimony from a Chinese witness, *ipso facto* incredible, had convicted him. The Court agreed, and cited the California legislature's Act Concerning Civil Cases, which prohibited the testimony of blacks or Indians against whites, as well as an 1850 criminal proceedings act which notoriously stated that "No Black, or Mulatto person, or Indian shall be allowed to give evidence in favor of, or against a white man."⁶⁰ California

⁵³ Ralph James Mooney, *Matthew Deady and the Federal Judicial Response to Racism in the Early West*, in *CHINESE IMMIGRANTS AND AMERICAN LAW* 241 (Charles McClain ed., 1994).

⁵⁴ Examples of habeas corpus cases may be found in *Immigration documents miscellany* on file with the Ethnic Studies Library, Univ. of California, Berkeley.

⁵⁵ CHRISTIAN G. FRITZ, *FEDERAL JUSTICE IN CALIFORNIA: THE COURT OF OGDEN HOFFMAN, 1851-1891* 213, 220 (1991).

⁵⁶ FRITZ, *supra* note 55, at 217, 223, 231, 233, 236, 241, 246-47, 248; LUCY E. SALYER, *LAW HARSH AS TIGERS: CHINESE IMMIGRANTS AND THE SHAPING OF MODERN IMMIGRATION LAW* 18-19, 69, 70, 72, 75, 92 (1995).

⁵⁷ McCLAIN, *IN SEARCH OF EQUALITY*, *supra* note 17, at 42.

⁵⁸ J.A.C. Grant, *Testimonial Exclusion Because of Race: A Chapter in the History of Intolerance in California*, 17 *UCLA Law R.* 193 (1969-70).

⁵⁹ *People v. Hall*, 4 Cal. 399 (Cal. 1854).

⁶⁰ *Id.* at 400.

Supreme Court Chief Justice Hugh C. Murray felt compelled to describe a lengthy, convoluted, and far-fetched rationale for why Chinese were co-equivalent with Indians, since Columbus had set sail for south Asia and thought he had arrived there in the West Indies of the Caribbean.⁶¹ Already openly racist, it seems (in retrospect) almost bizarre that Murray and the Court felt obliged to dress up their bigotry in the language of legal niceties; but so it went. As J.A.C. Grant pointed out, California's siding with the Union during the Civil War helped make the Chinese a substitute "devil to whip" in place of blacks. Regarding court testimony, in 1863 (halfway through the war) the state legislature amended its civil and criminal statutes to omit mention of blacks but to include "Mongolians, Chinese, or Indians."⁶²

The Reverend William Speer and other white elites denounced the *Hall* decision, as well as the 1870 *People v. Brady* decision that reiterated *Hall's* racism, despite imminent state legislative rectification, effective 1872.⁶³ After 1872, Chinese testimony experienced, at best, a mixed fate in and beyond California. Racists like U.S. Supreme Court Justice Stephen Field continued to condemn Chinese as somehow genetically dishonest.⁶⁴ Again in the minority, Judge Hoffman knew that Chinese witnesses could be dishonest, just like any other group of witnesses, but generally regarding them as competent witnesses.⁶⁵ In a more typical case, *Quock Ting v. United States* (1891), the majority decision dismissed Chinese testimony as unreliable,⁶⁶ in what amounted to an arbitrary and capricious decision. Judge John De Haven made at least one arbitrary and capricious dismissal of apparently consistent Chinese testimony (in *Woey Ho v. United States*, 1901), but more generally relied upon consistency or lack thereof in determining the reliability of Chinese witnesses, as did U.S. Commissioner of Immigration, E.H. Heacock.⁶⁷ On the other hand, collectors of customs in California regularly sought white witnesses to verify the citizenship of returning Chinese, which would have been detrimental were it not for the courts' and judges' counteracting policies.⁶⁸ Finally, Judge Maurice T. Dooling not only accepted Chinese testimony, but considered it no less than "ordinary fairness" for Chinese to be allowed to cross-examine witnesses against them in deportation cases.⁶⁹

Despite California's statutory rectification of disallowed Chinese testimony, a West Virginia court cited *Brady* in an 1877 case, claiming this

⁶¹ *Id.* at 401-03.

⁶² Grant, *supra* note 58, at 197.

⁶³ MCCLAIN, *In Search of Equality*, *supra* note 17, at 22, 36, 294, n.98; *People v. Brady* 40 Cal. 198 (Cal. 1870); Grant, *supra* note 58, at 201.

⁶⁴ FRITZ, *supra* note 55, at 237-38; Salyer, *supra* note 56, at 78.

⁶⁵ FRITZ, *supra* note 55, at 236, 241.

⁶⁶ *Quock Ting v. United States*, 140 U.S. 419 (1891).

⁶⁷ SALYER, *supra* note 56, at 79-80, 81.

⁶⁸ *Id.* at 83.

⁶⁹ *Id.* at 191.

meant racial discrimination in matters of testimony did not violate the Fourteenth Amendment.⁷⁰ In 1925, a Mississippi court claimed *Hall* justified their maintenance of white supremacy.⁷¹ One of the immediate and unfortunate reactions against *Hall* was the Chinese themselves making racist claims about their superiority over those of African and indigenous dissent.⁷² Although apparently not designed as such by the white majority, this effect was the equivalent to a “divide and conquer” strategy among racial and ethnic groups, at least for the time being.

III. UTILITARIAN FRIENDSHIPS DURING THE PROGRESSIVE ERA AND BEYOND

The Progressive Era in United States history spans the 1890s to the 1920s. Historians characterize this period as one containing multiple reform movements intent upon such diverse efforts as banning alcohol consumption, eradicating political corruption, limiting or banning child labor, improving sanitation in meat packing plants, and many more crusades. During the Progressive Era, University of Chicago especially became a center of multi-ethnic outreach, associated in part with the Settlement House movement design to help the urban poor, who were often immigrants.⁷³ University of Chicago Sociologist Robert Park in particular promoted a theory that immigrants would inevitably, if gradually, assimilate into American society.⁷⁴

For foreigners abroad and in the United States, the Progressive Era might have seemed like the best of times and worst of times. There was the multi-ethnic optimism of America as a haven for immigrants and native-born alike.⁷⁵ But this was also a period of intensifying nativism and immigrant resistance to or doubts about assimilation.⁷⁶ It was a time when ideas of Social Darwinism⁷⁷ and cultural absolutism⁷⁸ were on the rise, as well as

⁷⁰ *State v. Strauder* 11 W.Va. 745 (W.Va. 1877).

⁷¹ *Rice v. Gong Lum*, 139 Miss. 760 (Miss. 1925).

⁷² MCCLAIN, IN SEARCH OF EQUALITY, *supra* note 17, at 22. The racism of minorities remains hopelessly under-studied, perhaps mainly because of preoccupation with power differentials.

⁷³ MAUREEN A. FLANAGAN, AMERICA REFORMED: PROGRESSIVES AND PROGRESSIVISMS, 1890s-1920s, 35-39 (2007).

⁷⁴ STEPHEN E. CORNELL & DOUGLAS HARTMANN, ETHNICITY AND RACE: MAKING IDENTITIES IN A CHANGING WORLD 6, 43 (1998). Park led the gathering of information for the SURVEY OF RACE RELATIONS cited in this article. The collection is housed at the Hoover Institution Archives, Stanford Univ.

⁷⁵ CORNELL & HARTMANN, *supra* note 74, at 5, 6, 43; ALAN DAWLEY, STRUGGLES FOR JUSTICE: SOCIAL RESPONSIBILITY AND THE LIBERAL STATE 257 (1993); HIGHAM, *supra* note 11, at 133; MATTHEW F. JACOBSON, BARBARIAN VIRTUES: THE UNITED STATES ENCOUNTERS FOREIGN PEOPLES AT HOME AND ABROAD, 1876-1917 204-209 (2000).

⁷⁶ HIGHAM, *supra* note 11, at ch. 4; JACOBSON, *supra* note 75, at 209-213.

⁷⁷ HOFSTADTER, *supra* note 13, at 5; *see also* JACOBSON, *supra* note 75, at ch. 4.

⁷⁸ Dawley, *supra* note 75, at 263-64; HIGHAM, *supra* note 11, at 158, 165.

the beginnings of cultural relativism and the roots for modern cosmopolitanism.⁷⁹

Toward the end of the Progressive Era, the 1924 Johnson-Reed Act⁸⁰ created America's first quota system for immigration. This was the culminating triumph of years of nativist agitation⁸¹ and influenced the course for immigration policy until the mid-1960s.⁸² But the preceding decades had not been dominated by nativism alone. There were the usual surges of anti-immigrant sentiment associated with economic depression, especially 1893-97.⁸³ Still, the urban settlement house projects aimed primarily at immigrant workers⁸⁴ had mitigated, to some extent, anti-immigrant hostility.⁸⁵ On the foreign front, the United States boasted its first imperialist triumph with the 1898 conclusion of the Spanish-American War.⁸⁶ America gained control over the last major vestiges of Spain's old colonial empire, including the Philippines and Guam in the Pacific and Cuba and Puerto Rico in the Caribbean. Ironically, this victory brought both a renewed confidence in America's ability to absorb foreigners (the original "melting pot" concept)⁸⁷ and the imperialistic racism of the "White Man's Burden" variety.⁸⁸ The former represented the ongoing continental "mission" aspect⁸⁹ of the United States, while the latter reflected something of an Anglophile repetition of the "exclusive club" abroad.⁹⁰

The 1882 Chinese Exclusion Act⁹¹ did not pacify the racial nativists. 1885-1886 witnessed major property destruction and Chinese deaths in Rock Springs, Wyoming; Seattle and Tacoma, Washington; Los Angeles, and in Portland, Oregon.⁹² Extension and tightening of the 1882 Act arrived a decade later with passage of the 1892 Geary Act.⁹³ But violence and

⁷⁹ HIGHAM, *supra* note 11, at 22, 122, 125, 153, 304.

⁸⁰ Also called the National Origins Act of 1924. Pub. L. No. 68-139, 43 Stat. 153 (1924).

⁸¹ HIGHAM, *supra* note 11, at 186, 195, 204-12, 222-35, 261-63, 324-30.

⁸² ROGER DANIELS, *COMING TO AMERICA: A HISTORY OF IMMIGRATION AND ETHNICITY IN AMERICAN LIFE* 338-44 (2002).

⁸³ HIGHAM, *supra* note 11, at 68.

⁸⁴ FLANAGAN, *supra* note 73, at 35-39; HIGHAM, *supra* note 11, at 238-42.

⁸⁵ FLANAGAN, *supra* note 73, at 284; HIGHAM, *supra* note 11, at 117, 122.

⁸⁶ JACOBSON, *supra* note 75, at 263-65; HIGHAM, *supra* note 11, at 108-10, 144-45, 148; NELL I. PAINTER, *STANDING AT ARMAGEDDON: THE UNITED STATES 1877-1919*, 389 (1987).

⁸⁷ *Id.*

⁸⁸ *Id.*

⁸⁹ MIKISO HANE, *MODERN JAPAN: A HISTORICAL SURVEY* 400-401 (1986) (quoting Fosco Maraini in *THE JAPAN EXPERIENCE* 10-11 (Ronald Bell ed., (1973))). See also HIGHAM, *supra* note 11, at 22.

⁹⁰ This derived from the earlier, comparatively harmless "Cult of Anglo-Saxonism" that began to draw ominous racist overtones from 19th century German Romantic nationalism. See Fine & Cohen, *supra* note 5, at 145-146; HIGHAM, *supra* note 11, at 10, 132, 133, 134, 136-44; HOFSTADTER, *supra* note 13 at 172, 183; PAINTER, *supra* note 86, at 389-90.

⁹¹ Chinese Exclusion Act (May 6, 1882) 22 Stat. 58, c. 126.

⁹² McCLAIN, *IN SEARCH OF EQUALITY*, *supra* note 17, at 173-74.

⁹³ Geary Act, 27 Stat. 25, ch. 60 (1892).

hatred continued against Chinese, of course, even while racial nativist fear of an impending “invasion” began to diminish. These anti-immigration policies targeting the Chinese would not be reversed until 1943⁹⁴ when, belatedly, China’s alliance with the U.S. during World War II won such a minimal concession. In any case, with Chinese immigration effectively reduced by the late nineteenth century, racial nativist Americans found a new “Yellow Peril” in Japan.⁹⁵

Japan was unique in the history of imperialism.⁹⁶ Unlike all other targets of Western domination, both Japan and Ethiopia⁹⁷ resisted direct colonization during the age of imperialism. But only Japan actually became an imperial power itself.⁹⁸ Japan defeated China in the first Sino-Japanese War, 1894-95,⁹⁹ then triumphed in the Russo-Japanese War of 1904-05.¹⁰⁰ Japanese territorial expansion proceeded into Taiwan and the Korean peninsula, and by the 1930s into Manchuria and China itself.¹⁰¹ Some Americans grudgingly respected Japan for its meteoric modernization and military industrialization, but many more feared her.¹⁰² Japanese immigrants to the U.S. noticed an immediate rise in prejudice. Californians formed the Asiatic Exclusion League in 1905,¹⁰³ and passed Alien Land Laws targeting the Japanese in 1913¹⁰⁴ and 1920.¹⁰⁵ Mirroring earlier anti-Chinese legislation, the 1924 Johnson-Reed Act ended Japanese immigration altogether.¹⁰⁶

Seemingly paradoxically, continuing and deepening racial nativism in the U.S. between the 1890s and 1920s did not deter interracial friendship.

⁹⁴ This was the Magnuson Act, also known as the Chinese Exclusion Repeal Act of 1943, Act Dec. 17, 1943, ch. 344, § 1, 57 Stat. 600 (1943).

⁹⁵ DANIELS, *POLITICS OF PREJUDICE*, *supra* note 18, at ch.5; HIGHAM, *supra* note 11, at 166, 173-174; HOFSTADTER, *supra* note 13, at 189.

⁹⁶ For an interesting overview approach to Western imperialism, see IAN COPLAND, *THE BURDEN OF EMPIRE: PERSPECTIVES ON IMPERIALISM AND COLONIALISM* (1990).

⁹⁷ MICHAEL W. DOYLE, *EMPIRES* 218-20 (1986). Although Ethiopia successfully resisted domination during the 19th century era of imperialism, it nevertheless fell under Italian domination during the World War II era. See CHARLES S. MAIER, *AMONG EMPIRES: AMERICAN ASCENDANCY AND ITS PREDECESSORS* 39, 105 (2000).

⁹⁸ HANE, *supra* note 89, at chs. 5-9.

⁹⁹ *Id.* at 157-62. See also S. C. M. PAINE, *THE SINO-JAPANESE WAR OF 1894-1895: PERCEPTIONS, POWER, AND PRIMACY* (2003).

¹⁰⁰ HANE, *supra* note 89, at 171-79. See also *THE RUSSO-JAPANESE WAR IN GLOBAL PERSPECTIVE: WORLD WAR ZERO* (John W. Steinberg et al. eds., 2005).

¹⁰¹ HANE, *supra* note 89, at 253-58, ch. 13.

¹⁰² HOFSTADTER, *supra* note 13, at 189.

¹⁰³ DANIELS, *POLITICS OF PREJUDICE*, *supra* note 18, at 27.

¹⁰⁴ 1913 Cal. Stat. 206. See also Chuman, *supra* note 18, at 46-51; DANIELS, *ASIAN AMERICA*, *supra* note 18, at 145; DANIELS, *POLITICS OF PREJUDICE*, *supra* note 18, at 58-64; KIM, *supra* note 11, at 126.

¹⁰⁵ Initiative Act of 1920, 1921 Cal. Stat. 1921, lxxxiii. See also Chuman, *supra* note 18, at 76-80; DANIELS, *POLITICS OF PREJUDICE*, *supra* note 18, at 88-92; KIM, *supra* note 11, at 126.

¹⁰⁶ HIGHAM, *supra* note 11, at 324.

In many cases Asians and Caucasians seemed to grow more accustomed to one another and extended utilitarian friendships into friendships of pleasure and even the most profound sort of personal friendship. Again, it is the harsh historical context that makes these benevolent interracial relationships so striking.

In the western United States a great number of prominent Asian businessmen, most of them property owners, cultivated very congenial exchanges with white businessmen.¹⁰⁷ Here the situation tended to become more overtly political, as economic considerations motivated white businessmen and their organizations to speak out in defense of their Asian associates amidst an atmosphere of discrimination, racism, and exclusion.¹⁰⁸ Still, while political economy might primarily occupy the motivations of business people, they also typically developed friendships that transcended such practical considerations. George Shima (1864-1926), the millionaire California “potato king,”¹⁰⁹ Chin Lung (another “potato king”),¹¹⁰ Lee Bing, Seid Back Sr., Seid Chee (California, Oregon, and Washington state businessmen),¹¹¹ Minori Yasui (Hood River business and community leader),¹¹² Lung On of John Day, Oregon,¹¹³ and a great many others had extensive business ties and friendships among white people.¹¹⁴

¹⁰⁷ GUNTHER BARTH, BITTER STRENGTH: A HISTORY OF THE CHINESE IN THE UNITED STATES, 1850-1870 181 (1964); PARDEE LOWE, FATHER AND GLORIOUS DESCENDENT 75-78 (1943); SALYER, *supra* note 56, at 105; LIPING ZHU, “A CHINAMAN’S CHANCE” ON THE ROCKY MOUNTAIN MINING FRONTIER, IN CHINESE ON THE AMERICAN FRONTIER 160, 164-66 (2001).

¹⁰⁸ CHUMAN, *supra* note 18, at 129-30; SALYER, *supra* note 56, at 167. *See also Memorial of the Six Chinese Companies: An Address to the Senate and House of Representatives of the United States* 44th Cong. (1877); *The Other Side of the Chinese Question: Testimony of California’s Leading Citizens* 49th Cong. (1886).

¹⁰⁹ DANIELS, ASIAN AMERICA, *supra* note 18, at 134-35; DANIELS, COMING TO AMERICA, *supra* note 82, at 253-54.

¹¹⁰ SUCHENG CHAN, THIS BITTERSWEET SOIL 209, 212, 357 (1989).

¹¹¹ JOSEPH GASTON, PORTLAND, OREGON: ITS HISTORY AND BUILDERS, VOL. III 347-48 (1911); Portland City, OREGONIAN (June 6, 1936) at p. 21, c. 5; Jeff David Pricco, Homeless Strangers Among Us: The Chinese Community of Portland, Oregon, 1860-1900 59 (1978) (unpublished M.A. thesis, Univ. of Oregon) (on file with the Univ. of Oregon Libraries); Rose Marie Wong, Sweet Cakes, Long Journey: A Social and Urban History of Portland Oregon’s First Chinatown 122, 129-33, (1994) (unpublished Ph.D. dissertation, Univ. of Washington) (on file with the Univ. of Washington Libraries).

¹¹² LAUREN KESSLER, STUBBORN TWIG: THREE GENERATIONS IN THE LIFE OF A JAPANESE AMERICAN FAMILY 28-33, 51, 70 (1993).

¹¹³ Canyon City, Oregon, BLUE MOUNTAIN EAGLE, June 17, 1927, at 1, c. 4; Canyon City, Oregon, BLUE MOUNTAIN EAGLE, Dec. 31, 1937, at 8, c.1-2; Canyon City, Oregon, BLUE MOUNTAIN EAGLE, Dec. 20, 1940, at 8, c. 4-5; Chia-lin Chen, A GOLD DREAM IN THE BLUE MOUNTAINS: A STUDY OF THE CHINESE IMMIGRANTS IN THE JOHN DAY AREA [OF] OREGON, 1870-1910, 120 (1972) (unpublished M.A. thesis, Portland State Univ.) (on file with the Branford P. Millar Library). *See also* Will Sarvis, *Gifted Healer: Ing Hay and the Chinese Medical Tradition in Eastern Oregon, 1888-1948*, 44:3 J. OF THE WEST 62 (2005).

Asian businessmen in general had much-lauded reputations for honesty and responsibility, which their white counterparts cherished all the more in an era of high speculation, volatile capitalism, bankruptcy, and cancellation of debts.¹¹⁵ Many whites reported losses with their white counterparts but, conspicuously, none with their Asian counterparts. A Caucasian president of an Oregon fruit co-op wrote to Masuo Yasui, "For a number of years we have taken quite an interest in financing various Japanese who have gone into the strawberry and fruit business in Hood River Valley and have done so largely because they have been connected with your firm . . ."¹¹⁶

Much more common than interaction in elite business circles were labor-employer relations. Long traditions of an industrious work ethic permeate East Asian cultures, which was bound to appeal to many white observers in America, steeped in their own Protestant work ethic.¹¹⁷ As early as 1884 the Milwaukee (Oregon) *Sentinel's* editor (Horace Rablee) acknowledged the efficacy of Chinese labor, citing industriousness, inexpensiveness, neatness, lack of complaining, and rapid learning on the part of the Chinese.¹¹⁸ The *Anacortes American*, a newspaper in the small coastal town of Anacortes, Washington, was usually filled with racist editorials and reports against Asians. Nevertheless, it reported a conversational exchange in 1900 that illustrated the grudging respect some racist whites held for industrious Asian workers. "You people always have a lot of Chinaman [sic] working here," remarked a Hotel Taylor guest, visiting from Whatcom County to the north. "And at Whatcom you've got a lot of loafing paupers who don't work," rejoined his companion.¹¹⁹

Before the 1924 Johnson-Reed Act, the *Anacortes American* and a great many other newspapers along the Pacific Coast generally expressed alarm at the influx of Asian workers. With Chinese immigration severely limited after 1882, racists began targeting the Japanese.¹²⁰ Fear over Japanese labor competition, apparently inflamed by American labor agitators, contributed to rioting in Vancouver, British Columbia, in September

¹¹⁴ See also the following interviews in the SURVEY OF RACE RELATIONS Collection: Mar Sui Haw Interview (box 28, file 244) at 3-5, 7; Andrew Kan Interview (box 27, file 178) at 10; Lew Kay Interview (box 27, file 194) at 4; J.S. Look Interview, *supra* note 49, at 6.

¹¹⁵ *Memorial of the Six Chinese Companies: An Address to the Senate and House of Representatives of the United States* 44th Cong. (1877); *The Other Side of the Chinese Question: Testimony of California's Leading Citizens* 49th Cong. (1886).

¹¹⁶ FRANK DAVEY, REPORT ON THE JAPANESE SITUATION IN OREGON 6-10 (1920); KESSLER, *supra* note 101, at 56-57. See also *supra* notes 38 and 39.

¹¹⁷ A prominent example of this can be found in the Rev. William Speer's observation, from a Protestant point of view. See SPEER, *supra* note 32, at 494, 495, 526, 529, 530, 582-83.

¹¹⁸ Portland City, OREGONIAN, July 22, 1884, at 4.

¹¹⁹ ANACORTES AMERICAN, Mar. 29, 1900, at 3.

¹²⁰ Examples include ANACORTES AMERICAN, Apr. 19, 1900, at 2; Apr. 26, 1900 at 2; Jan. 10, 1901, at 2; Jan. 12, 1905, at 4; Aug. 15, 1907; at 5; July 22, 1915, at 1.

1907.¹²¹ The American Federation of Labor (AFL) had previously joined nativists in opposing immigrant labor, though their focus had been the eastern United States and European immigration. During the 1893-97 depression, AFL president Samuel Gompers (an English Jew) had opposed the influx of Russian Jews into the cigar making industry. By the early 1920s, the AFL joined the clamor for immigration restriction that led to the 1924 quota law.¹²² So it was consistent for the AFL, in 1908, to oppose Asian labor, though they oddly specified the increasingly anachronistic “threat of Chinese labor.”¹²³

Still, there were remarkable exceptions to this labor competition and Caucasian disparaging of Asian and Asian-American labor. An anonymous Japanese servant published his employment experiences in the *New York Independent* on September 21, 1905.¹²⁴ This man’s employers ranged from intolerable to very friendly. One of his favorite employers was a Princeton graduate and his wife, both of whom treated him very kindly.¹²⁵ The servant worked for them almost three years while he finished high school, and only the absence of a local college or university compelled him to depart, which he wrote he “exceedingly” hated to do. His next employer was a haughty, dictatorial woman, and he left her employ almost immediately.¹²⁶ While in college this servant’s best position was as a cabin boy on a yacht, where he worked for a summer. He described the yacht’s owner and his wife as “very agreeable” and their children as “lovely and good-natured.” Unlike the dictatorial woman, who had made the social distance between them plain, the yacht people included their servant in conversations and helped him with his education expenses even after he had left their employ.¹²⁷

Similar to elite business relationships, true affection could also develop among non-elite Asian workers and the various whites who hired them. This was especially likely in domestic environments, when whites frequently described their Asian cooks or domestics as virtually “one of the family.”¹²⁸ Perhaps no clearer example of this was reflected in the life of Luke

¹²¹ Masako Iino, *Japan’s Reaction to the Vancouver Riot of 1907*, 60 BR. COLUM. STUDIES 60 28 (Winter 1983-84); Howard H. Sugimoto, *The Vancouver Riot and Its International Significance*, 64:4 PAC. NORTHWEST Q. 163 (1973).

¹²² HIGHAM, *supra* note 11, at 71, 305-06.

¹²³ SAMUEL GOMPERS AND HERMAN GUTSTADT, SOME REASONS FOR CHINESE EXCLUSION, MEAT VS. RICE: AMERICAN MANHOOD AGAINST ASIATIC COOLIEISM, WHICH SHALL SURVIVE? (1908).

¹²⁴ This article is reproduced in SURVEY OF RACE RELATIONS, box 26, file 109.

¹²⁵ *Id.* at 7-8.

¹²⁶ *Id.* at 8-9.

¹²⁷ *Id.* at 9-12.

¹²⁸ Wong Chin Foo, *The Chinese in New York*, COSMOPOLITAN, June 1888, at 300; EVELYN NAKANO GLENN, ISSEI, NISEI, WAR BRIDE: THREE GENERATIONS OF JAPANESE AMERICAN WOMEN IN DOMESTIC SERVICE 154-158 (1986); HAMILTON HOLT, THE LIVES OF UNDISTINGUISHED AMERICANS: AS TOLD BY THEMSELVES 166-171 (1906); HERBERT JOHNSON, DISCRIMINATION AGAINST THE JAPANESE IN CALIFORNIA 41-43 (1907); Evelyn

Chess. Luke Chess, born in 1890 to Chinese parents in San Francisco, moved to Genoa, Nebraska in 1910 to work for Mrs. Ed L. Burke of the Kent & Burke Cattle Ranch. In 1924 in his faltering command of English he wrote, “[S]he was the most kind hearted lady to me in my life, I took her as a mother to me, and her’s care take and love to me as to her own son [S]he not only known of my happiness, but my troubles and sorrow as well, until to to [sic] dated she is my best friend I have had.”¹²⁹

Obviously white employers could also be condescending and racist, but this should not discount other instances genuine good will.¹³⁰ In fact, genuine affection could sometimes be intermixed with the unfortunate bigotry and hierarchical cultural absolutism so typical of the era that failed to appreciate Asian culture in its own right. But what seems ambiguous or dubious now should not be imposed upon the past lest we commit the historian’s fallacy of presentism.¹³¹ In regard to Ye Gon Lun, a Chinese boy in California during the late nineteenth century, Ira M. Condit wrote,

He came into the home of Honorable N. Greene Curtis, of Sacramento, a little boy of only nine years of age, and fresh from his heathen home in China. He was meant to be only a servant boy. By the loveliness of his character, and his wonderful faithfulness to duty, he soon won his way into the hearts of this household. They learned to feel towards him and to treat him more as a son than as a servant. In natural uprightness and nobleness of nature he was far above the average of his countrymen, and soon became separated in sympathy from them, so that he scarcely seemed to be Chinese at all.¹³²

But there were other qualifications of the paternal-symbiotic situation, more positive even from a Western cultural perspective. Contrary to stereotypes of docility, dissatisfied Asian workers often left their individual jobs or sometimes collectively rebelled against employers through strikes and boycotts.¹³³ The family of John Reed, famous author of *Ten Days that*

Sibley Lampman, *As It Was*, OREGON HIST. Q., Dec. 1975, at 369, 371-72; ROSE LEE, GROWTH AND DECLINE OF CHINESE 218-19 (1978); MITZI ASAI LOFTUS, MADE IN JAPAN AND SETTLED IN OREGON 18-19 (1990); City of Portland, OREGONIAN, July 22, 1884, at 4; ROSSITER W. RAYMOND, STATISTICS OF MINES AND MINING IN THE STATES AND TERRITORIES WEST OF THE ROCKY MOUNTAINS 3-6 (1872); SHIH-SHAN HENRY TSAI, *supra* note 23, at 58.

¹²⁹ *Autobiography of Mr. Luke Chess*, Oct. 1924, SURVEY OF RACE RELATIONS, box 30, file 18, at 5.

¹³⁰ Paul Merriam dismisses such good will as a situation in which elite whites who were not competing in the work force valued the high quality of Chinese service. Obviously a good deal of truth must lie in this depiction, but the qualification of its over-simplicity presents a microcosm of this essay’s main thesis. Paul G. Merriam, *Portland, Oregon, 1840-1890: A Social and Economic History*, 71 (Univ. of Oregon, PhD dissertation, 1971).

¹³¹ DAVID HACKETT FISCHER, HISTORIANS’ FALLACIES 135-40 (1970).

¹³² IRA M. CONDIT, THE CHINAMAN AS WE SEE HIM 181 (1900).

¹³³ CHAN, THIS BITTERSWEET SOIL, *supra* note 110, at 332-33; GLENN, *supra* note 128, at 161-64; LEE MERIWETHER, THE TRAMP AT HOME 175-77, 191 (1889). *See also*

Shook the World, employed a Chinese domestic who was also a proud property owner, and thus hardly servile.¹³⁴ Evelyn Nakano Glenn, who described many good relations between Japanese servants and their employers, also described servants who tried (with varying success) to dictate the terms of their employment and, if necessary, resigned rather than tolerate demeaning or otherwise unhappy working conditions.¹³⁵ Ironically, some of today's purported proponents of interracial goodwill often ironically make cross-cultural mistakes, and inadvertently impose their values upon others. Regarding historic Asian-Caucasian friendships, American critics of paternalism sometimes fall directly into this trap. It is worth a more deep examination. A Chinese ranch worker will illustrate the point.

Ah Sam, a ranch hand for the Jewett family in California before World War II, insisted on removing himself to Chinatown after the Jewetts sold their ranch.¹³⁶ Fiercely loyal to the family, Ah Sam refused to work for another, not even Hugh Jewett's cousin. Hugh Jewett apparently failed to convey the fact that his own family was going out of the ranching business, and remained concerned for Ah Sam's future and well being.¹³⁷ Finally, it occurred to Hugh Jewett to appeal to Ah Sam's very sense of family loyalty. He informed Ah Sam that his cousin was ill, in dire need of his help, and that Ah Sam might consider the cousin a member of the extended Jewett family. The result was a new job for Ah Sam and the delivery of much needed high quality labor.¹³⁸

Probably without realizing it, Jewett had appealed to Ah Sam's Confucian sense of loyalty and constancy in friendship. From an Aristotelian point of view this was mainly a utilitarian sort of friendship,¹³⁹ but a Confucian perspective strongly suggests a much deeper sense of dedication than mere utilitarianism.¹⁴⁰ Beyond all particular cultural considerations, however, this relationship demonstrated an almost fierce element of good will.

As mentioned, scholars have criticized the paternalism that sometimes characterized these types of employer-laborer relationships, but this presents a potential Anglo-American cultural distortion and thus what David Hackett Fischer described as the "fallacy of ethnomorphism."¹⁴¹ Anglo-Americans, with their long and colorful heritage of hyper-individualism,

Anonymous Japanese Servant, N.Y. INDEPENDENT, Sept. 21, 1905, SURVEY OF RACE RELATIONS, box 26, file 109.

¹³⁴ Fred DeWolfe, Ed., *Portlander John Reed Remembers Lee Sing, His Family's Chinese Servant*, 97:3 OR. HIST. Q. 365 (1996).

¹³⁵ GLENN, *supra* note 128, at 160-64.

¹³⁶ "Ah Sam" (or "Yellow Sam"): from Hugh S. Jewett reminiscences (Call# BANC MSS 80/84 c: fol. 6, Bancroft Library, Univ. of California-Berkeley).

¹³⁷ *Id.*

¹³⁸ *Id.* See also CHAN, THIS BITTERSWEET SOIL, *supra* note 110, at 339-40; TSAI, *supra* note 23, at 35.

¹³⁹ ARISTOTLE, NICOMACHEAN ETHICS at 214 (Broadie & Rowe trans.) (c. 384 B.C.E.).

¹⁴⁰ CONFUCIUS, ANALECTS OF CONFUCIUS bks. V:25 and XV:5 at 114, 194 (Waley, trans.) (c. 206 B.C.E.).

¹⁴¹ FISCHER, *supra* note 131, at 224-25.

tend to disparage paternalism uniformly.¹⁴² East Asians, with even longer traditions of group-orientation, champion loyalty, honesty, diligence, and other forms of integrity in highly social relationships — all consistent with Confucian ideas of harmonious and virtuous social relationships.¹⁴³ Being dependent upon an employer, from an Asian perspective, was not necessarily dishonorable.¹⁴⁴ The noted scholar Sucheng Chan, in fact, does not necessarily describe these sorts of relationships as “paternal” at all, but rather as symbiotic.¹⁴⁵ Sucheng Chan makes an excellent point. There was good reason for recognizing these relationships outside of western notions of paternalism, even if symbiotic relationships were not always benevolent.

Mr. Preble of the Curtis Packing Company of Long Beach, California, highly praised their Japanese workers for their cleanliness, hard work, trustworthiness, cheerfulness, and high levels of education — but noted that they strongly resisted being “bossed” by whites, in contrast to complying with the orders of Japanese supervisors.¹⁴⁶

Monica Sone’s father, a Japanese hotel owner in pre-war Seattle, was “paternalistic” himself toward his Caucasian employees, who were “like family” to the Sones. For example, Joe Subotich, a white man down on his luck, experienced a change in fortune by meeting Mr. Sone. “Father trusted him,” his daughter remembered, “and Joe’s gratitude knew no bounds. Eventually Joe became our night watchman and Father’s loyal friend.”¹⁴⁷ Thus it remains important to avoid reading Anglo-American cultural preferences for individualism into pre-war Asian-Caucasian employee-worker

¹⁴² Leverenz goes on to describe paternalism as more complicated and ambiguous than its detractors. DAVID LEVERENZ, *PATERNALISM INCORPORATED: FABLES OF AMERICAN FATHERHOOD, 1865-1940* 1 (2003). *See also* ANDREA TONE, *THE BUSINESS OF BENEVOLENCE: INDUSTRIAL PATERNALISM IN PROGRESSIVE AMERICA* (1997). *See generally* BRANNON COSTELLO, *PLANTATION AIRS: RACIAL PATERNALISM AND THE TRANSFORMATIONS OF CLASS IN SOUTHERN FICTION, 1945-1971* (2007); MELTON A. MCLAURIN, *PATERNALISM AND PROTEST: SOUTHERN COTTON MILL WORKERS AND ORGANIZED LABOR, 1875-1905* (1971); JOHN DAVID SMITH, ED., *RACIST SOUTHERN PATERNALISM* (1993).

¹⁴³ CONFUCIUS, *supra* note 140, at 85, 86-87, 114, 170, 194, 205.

¹⁴⁴ The usual stereotype depicts Western culture as individualistic and East Asian culture as group-oriented. There may be more truth in this than the average stereotype, but of course this simplicity fades upon further examination. For the origins of individualism in Western culture, *see generally* COLIN MORRIS, *THE DISCOVERY OF THE INDIVIDUAL, 1050-1200* (1972). For Western philosophical ideas about the individual from the 17th century to modern times, *see generally* BRIAN MORRIS, *WESTERN CONCEPTIONS OF THE INDIVIDUAL* (1991). For ambiguities about the dichotomy of Western individualism and Eastern collectivism, *see generally* JOHN D. GREENWOOD ET AL., *INDIVIDUALISM AND COLLECTIVISM IN MORAL AND SOCIAL THOUGHT IN THE MORAL CIRCLE AND THE SELF: CHINESE AND WESTERN APPROACHES* 163-73 (2003).

¹⁴⁵ CHAN, *THIS BITTERSWEET SOIL*, *supra* note 110, at 176-78, 185, 187, 189, 243, 361-64, 368.

¹⁴⁶ William Kinchelve, *A Study of the Japanese Fishermen at Curtis Packing Company in Long Beach*, SURVEY OF RACE RELATIONS, box 30, file 30, at 1-5.

¹⁴⁷ MONICA SONE, NISEI DAUGHTER 30-33 (1979).

exchanges, even while recognizing the crucial aspects of potential and actual exploitation.

Asian salmon cannery workers, timber mill workers, agricultural laborers, domestic servants, farm cooks, and other workers all participated in various symbiotic-paternal situations with their employers. While clearly a utilitarian type of friendship, white employers repeatedly and consistently sung the virtues of their Asian workers during the entire late nineteenth and early twentieth century period.¹⁴⁸ Again, such Caucasian attitudes stood in marked contrast to the dominant racism of the period that contained no shortage of anti-Asian labor agitation among working class whites.

Hundreds of Japanese worked in the timber mill industries of Washington state during the 1920s. A certain Mr. Ninemire, president of N & M Lumber Company in Rochester, Washington, expressed the typical praise of his Japanese workers. In a 1924 interview he described them as more dependable, more consistent, and harder working than whites.¹⁴⁹ R. Ode, Japanese foreman of the Eatonville Lumber Company, described how his Japanese mill workers were in high demand for moonlighting positions as carpenters. Their reputation as diligent workers came from their regular employment at the mill. "All my boys have been here for a long time and they all work like hell," Ode said in 1924.¹⁵⁰ A great many similar descriptions from the 1920s can be found at numerous other Washington mills such as Carlisle Pennell Lumber Company (Onalaska), the Ernest Dolge mill (Tacoma), St. Paul and Tacoma Lumber Company (Tacoma), Grays Harbor Commercial (Cosmopolis), the Doty Lumber and Shingle Company (Doty), and many others.¹⁵¹ While clearly a mutually utilitarian relationship, these mills also experienced wider varieties upon the theme, in-

¹⁴⁸ See Pricco, *Homeless Strangers*, CALIFORNIA HISTORICAL SOCIETY 57 (1886); C.H. Burnett, *Interview and Social Document of George T. Meyers*, July 27, 1924, SURVEY OF RACE RELATIONS, box 27, file 185, at 1-2; Marjorie McComber, *A Japanese Agricultural Community Near Fullerton, California*, May 26, 1924, SURVEY OF RACE RELATIONS, box 26, file 118, at 3-4.

¹⁴⁹ R.L. Olson, *Interview with Mr. Ninemire, President of the Company*, July 22, 1924, SURVEY OF RACE RELATIONS, box 28, file 210, at 3, 4.

¹⁵⁰ Ronald L. Olson, *Impressions, Attitudes and Incidents*, SURVEY OF RACE RELATIONS, box 28, file 204, at 23.

¹⁵¹ Ronald L. Olson, *The Orientals in the Lumber Industry in the State of Washington*, SURVEY OF RACE RELATIONS, box 27, file 198, at 9, 10, 12, 15; A.R. Johnson *Interview*, SURVEY OF RACE RELATIONS, box 27, file 200, at 4; *Jimmy Harwell Interview and Mr. Patterson Interview*, SURVEY OF RACE RELATIONS, box 28, file 203, at 6, 9; Olson, *Impressions*, *supra* note 150, at 8, 10, 21; *Mr. Stout Interview*, *E.S. Bennett Interview*, *White Mill Truck Driver Interview*, SURVEY OF RACE RELATIONS, box 28, file 207, at 3, 5, 7; *Mr. Ellis Interview*, SURVEY OF RACE RELATIONS, box 28, file 201, at 5; *Mr. Matsui Interview*, *Mr. Holmes Interview*, *Mr. Rogers Interview*, *E.G. Griggs to Ronald L. Olsen*, August 19, 1924, SURVEY OF RACE RELATIONS, box 28, file 202, at 4, 8, 9, 12, 13; R.L. Olson, *Doty Lumber and Shingle Company*, SURVEY OF RACE RELATIONS, box 28, file 206, at 1, 6.

cluding exploitation of harder working Japanese, but also instances of interracial horseplay and even “true friendship.”¹⁵²

A certain Mr. Kier fished and bought fish in southern California for more than twenty years during the early twentieth century.¹⁵³ Of the multiple racial and ethnic groups in the area, Kier singled out the Japanese fisherman as the best. Kier admired their educated approach to their working class endeavor, wherein the Japanese studied weather and ocean conditions and how they affected various fish species.¹⁵⁴ Kier praised the Japanese for maintaining their boats and equipment better than other fishers, for the unequaled quality of fish they sold, and for their good will. “They have always been my friends, they always have a smile for you,” Kier said around 1924.¹⁵⁵

In some cases a delightful cultural interaction took place — such as the United States-Japan exchange that the Ike family organized for the McCombers in 1923.¹⁵⁶ The Ikes were Japanese tenant farmers on the McComber farm near Fullerton, California. They arranged for the McCombers to be received by various tourist bureau and chamber of commerce delegations when the latter visited Japan.¹⁵⁷ The Ikes performed this kindness for the McCombers just before California’s 1920 Alien Land Law took effect, excluding Japanese from leasing agricultural land.¹⁵⁸ It was a gesture of good will toward individual friends within a society becoming, at that time, increasingly discriminatory toward Japanese.¹⁵⁹

In Oakland, California, and Vancouver, British Columbia, there was apparently an easing of white-Asian racial tensions during the 1930s that may have indicated a wider pattern.¹⁶⁰ Why tensions would ease when economic depression increased competition for jobs remains unknown, but would seem to qualify to usual depiction of such conflict as a simple ex-

¹⁵² Olson, *Orientalism in the Lumber Industry*, *supra* note 151, at 15.

¹⁵³ Kinchelve, *supra* note 146, at 5-7.

¹⁵⁴ *Id.*

¹⁵⁵ *Id.*

¹⁵⁶ McComber, *A Japanese Agricultural Community*, *supra* note 148, at 5.

¹⁵⁷ *Id.*

¹⁵⁸ *Id.*

¹⁵⁹ Incidentally, countless whites helped Asians circumvent alien land laws of the 1910s and 1920s by purchasing land in their own names and holding acreage for Chinese and Japanese. Such an arrangement required a great deal of trust. If Caucasians honored their side of the agreement, the arrangement resulted in stronger bonds of friendship. Betrayal, of course, was irredeemable, without legal recourse. In all their outcomes, such arrangements remained by necessity secret. See KAZUO ITO, ISSEI: A HISTORY OF JAPANESE IMMIGRANTS IN NORTH AMERICA 663-665 (1973); Amy K. Buck, *Alien Land Laws: The Curtailing of Japanese Agricultural Pursuits in Oregon*, 49-51 (MA thesis, Portland State Univ., 1999).

¹⁶⁰ KAY J. ANDERSON, VANCOUVER’S CHINATOWN: RACIAL DISCOURSE IN CANADA, 1875-1980, at 144-177 (1991); L. EVE ARMENTROUT MA, HOMETOWN CHINATOWN: THE HISTORY OF OAKLAND’S CHINESE COMMUNITY 87-108 (2000).

planation for earlier conflicts, especially between Chinese and Irish immigrants.¹⁶¹

Beyond the workplace, Asians continued to find benevolent allies in the legal and political realms.¹⁶² In 1915, San Francisco Mayor James Rolph, Jr. (1912-1931) used his political power directly to subvert discrimination.¹⁶³ The United States government had invited about twenty Chinese men to participate in the 1915 Panama-Pacific International Exposition in San Francisco. Rolph intervened with immigration authorities who refused to admit some of these men on the grounds of purported hookworm and trachoma infection — but apparently were actually detaining them aboard ship in order to extort them for expensive “medical treatment” of said maladies.¹⁶⁴ Rolph, who also served as president of the Exposition, pressured the immigration authorities to release the Chinese under his care and the Chinese participated as originally planned. David Young, a leader among the Chinese, remembered Rolph for his positive intervention and said Rolph also secured housing for the Chinese after the Exposition. Young attributed American prejudice to the laboring class and European immigrants, but clearly appreciated Rolph’s judicious exercise of political power.¹⁶⁵

The Chinese Six Companies regularly retained attorneys to handle cases, which often became the specialties of certain lawyers.¹⁶⁶ Unlike the legal fight for civil rights in the American South, there was often a great deal of money to be made in lawsuits affecting Asians.¹⁶⁷ Various ethnic

¹⁶¹ Much of this is associated with labor competition, epitomized by Dennis Kearney (1847-1907), co-organizer of the Workingman’s Party whose late 1870s anti-Chinese agitation helped instigate the 1882 Chinese Exclusion Act. See MCCLAIN, IN SEARCH OF EQUALITY, *supra* note 17, at 79-80, 309. See also HAMILTON HOLT, LIVES OF UNDISTINGUISHED AMERICANS 184-185 (1906); *Memorandum, Their Excellencies Pao and Li to the Commissioners*, 47th Cong., 2554 (Apr. 4, 1882); *The Chinese Panic*, HARPER’S WEEKLY, 306-307 (May 20, 1882). Several years before Kearney came on the scene, Mark Twain and the Rev. William Speer both blamed (“low class”) whites for anti-Chinese activity and sentiment, see generally MARK TWAIN, ROUGHING IT 391-392 (1872); SPEER, *supra* note 32, at 71, 530 (1870). For early 20th century accounts blaming the Irish and other European ethnic immigrants for anti-Asian racism, see generally *Mr. Banton and Anonymous White Engineer Interviews*, SURVEY OF RACE RELATIONS, box 28, file 208.

¹⁶² See also DANIEL JOHNSON, ANTI-JAPANESE LEGISLATION IN OREGON, 1917-1923, at 196-197 (1996); SALYER, *supra* note 56, at 62, 66 (1995); Robert R. Swartout, Jr., *In Defense of the West’s Chinese: Denny’s Brief for Li Hung-chang*, 83:1 OREGON HIST. Q. 27-29 (1982).

¹⁶³ *Life History and Social Document of David Young*, Aug. 29, 1924, SURVEY OF RACE RELATIONS, box 29, file 22, at 3-5.

¹⁶⁴ *Id.*

¹⁶⁵ *Id.*

¹⁶⁶ SALYER, *supra* note 56, at 40.

¹⁶⁷ *Id.* at 70-71.

organizations gathered dues to pool for lawsuits.¹⁶⁸ Churches and philanthropists also donated money.¹⁶⁹ In some circumstances attorneys must have enjoyed a lucrative trade, and therefore an interest in profit obscures their motives. In fact, one San Francisco lawyer (John Henry Boalt) was openly anti-Chinese, but suspended his personal beliefs to exploit pro-Chinese lawsuits.¹⁷⁰ In other cases, attorneys were on a retainer's fixed annual salary and were obliged to handle all cases, no matter how numerous, during that period.¹⁷¹ Some Caucasians and Asians simply remembered each other as honest businessmen and clients who had enjoyed an honest and fair exchange.¹⁷² Some attorneys likely had mixed motives involving economy and philanthropy. But some clearly came to have Asians' best interests at heart.

Every attorney, it should be remembered, faced potential public opposition for defending Asians at all, whether for profit, altruism, or mixed motives.¹⁷³ Louis Guernsey, a Los Angeles attorney with political ambition, publicly denounced his colleague, Robert Young, for representing local Japanese.¹⁷⁴ In fact, Guernsey claimed Young should be disbarred on these grounds.¹⁷⁵ Masuo Yasui, the Hood River leader, developed a very close friendship with Ernest C. Smith, a local lawyer.¹⁷⁶ Smith handled a great deal of legal work for the Hood River Japanese, but also invited Yasui to his house, served him tea, and suffered malicious backbiting and loss of business from area whites for his trouble.¹⁷⁷

¹⁶⁸ *Id.* at 47. Franklin B. Worley functioned as contract attorney for the Chinese League of Justice of America during the early 1900s, as did Carroll Cook for the Chinese Consolidated Benevolent Association (the "Six Companies") during the same period. See Carroll Cook correspondence (call# AAS ARC 2000/42, folders 2 and 14), and Chinese League of Justice of America miscellany (call# AAS ARC 2000/43) both Ethnic Studies Library, Univ. of California-Berkeley.

¹⁶⁹ CHUMAN, *supra* note 18, at 12-13.

¹⁷⁰ FRITZ, *supra* note 55, at 212.

¹⁷¹ MCCLAIN, *In Search of Equality*, *supra* note 17, at 336, 345-46.

¹⁷² Chin Cheung, leading Seattle merchant of the 1920s, spoke well of his attorneys, Walter Fulton and a certain Mr. Norman, see *Chin Cheung Interview*, SURVEY OF RACE RELATIONS, box 27, file 187, at 3. J.A. Russel, 35-year veteran criminal lawyer for the Seattle Chinese by 1924, described his clients as "very loyal and very honest," which, of course, were typical Confucian virtues, see *J.A. Russell Interview*, SURVEY OF RACE RELATIONS, box 24, file 27, at 2.

¹⁷³ FRITZ, *supra* note 55, at 212; MCCLAIN, *IN SEARCH OF EQUALITY*, *supra* note 17, at 64.

¹⁷⁴ Robert Young (Series A, #163; Series A, #306, #309, *Smith Collection*, available at Univ. of Oregon Special Collections).

¹⁷⁵ *Id.*

¹⁷⁶ KESSLER, *supra* note 112, at 61-62 (2008).

¹⁷⁷ *Id.* An anonymous critic of this paper for the *Pacific Historical Review* dismissed the idea of Masuo Yasui having any Caucasian friends because he eventually committed suicide. Aside from the ludicrous implication that friendlessness causes suicide or that people who commit suicide have no friends, this criticism also advances an inexcusably (and ironically) racist non-appreciation for Japanese culture, wherein suicide is sometimes the ultimate expression of honor. The place of suicide in Japanese culture has deep roots

In early 1923, San Francisco District Attorney Matthew Brady (in office 1919-1943) petitioned Angel Island's Commissioner of Immigration (a Mr. White) not to deport certain Chinese slave women after they had testified against their captors.¹⁷⁸ Brady, on the public payroll, likely had no other motive than good will. He accurately saw the women as victims and, after receiving their testimony, he wanted them paroled safely to Donaldina Cameron's much noted mission house.¹⁷⁹ Cameron (1869-1968), for all her flaws, was famous for going to great lengths to rescue Chinese slave women who were regularly kidnapped into prostitution.¹⁸⁰

Fred H. Lysons, 25-year veteran attorney for the Chinese in Seattle by 1924, actually became something of an amateur Sinologist. He first traveled to China in 1914 and ended up possessing an atypical appreciation for Chinese cultural phenomenon such as ancestor worship and *feng shui* (sometimes called Chinese geomancy).¹⁸¹ Aside from his legal work, then, perhaps Lysons might be credited with an interest in Asian culture, itself an anti-racist attitude.

While practically all benevolent pre-war legal affairs involved Asian clients and Caucasian attorneys and judges, the field of medicine exhibited something of a reversal in roles. Here Chinese doctors brought a centuries-old healing tradition so radically different from Western medicine as to attract an entire white clientele dissatisfied with orthodox practices. After all, orthodox Western medicine had recently sanctioned practices such as bleeding and purging, the latter definitely potentially harmful to ailing patients.¹⁸²

Chinese doctors had to endure inaccurate comparisons with late nineteenth and early twentieth century unorthodox healers, some of whom were the legendary "snake oil salesmen" charlatans of the American West.¹⁸³ Newspaper writers did not hesitate to equate Chinese physicians

in the centuries-old Samurai ethic and Zen Buddhism and is practically incomprehensible from a biased Western perspective with its notions of "suicide as sin." This misconstrued criticism illustrates the "fallacy of ethnomorphism;" see FISCHER, *supra* note 131, at 224-25.

¹⁷⁸ SAN FRANCISCO DAILY NEWS, Feb. 8, 1923.

¹⁷⁹ *Id.*

¹⁸⁰ For accounts on slave women, see generally SURVEY OF RACE RELATIONS, box 25, files 55 and 83. For an interview with Cameron, see SURVEY OF RACE RELATIONS, box 26, file 129. Donaldina Cameron's work in San Francisco was arguably the most famous among Christian benevolent efforts. But behind Cameron's supposed benevolence lay a great deal of dogmatism, racism, intolerance and other negative features, see also Laurene Wu McClain, *Donaldina Cameron: A Reappraisal*, 27:3 PAC. HISTORIAN 24-35 (1983).

¹⁸¹ *Fred H. Lysons Interview*, SURVEY OF RACE RELATIONS, box 27, file 192, at 1-4.

¹⁸² Gevitz provides a highly informative work distinguishing what might be described as the "loose dichotomy" between conventional and unconventional Western medicine of the era, see NORMAN GEVITZ, OTHER HEALERS: UNORTHODOX MEDICINE IN AMERICA 17-18 (1988).

¹⁸³ *Id.* at 6; Haiming Lu, *The Resilience of Ethnic Culture: Chinese Herbalists in the American Medical Profession*, 1:2 J. ASIAN AM. STUDIES 179, 180, 183 (1998).

with quacks, while others were grudgingly respectful.¹⁸⁴ Clearly all sorts of Western unorthodox practitioners paid for regular newspaper advertising and did not suffer as much proportional deprecation.¹⁸⁵ Furthermore, police often extorted Chinese doctors for practicing medicine without licenses, while health boards opposed their practice.¹⁸⁶ Chinese physicians also probably endured an inordinate number of terminally ill patients who came to them only as a last resort.¹⁸⁷ Nevertheless, successful Chinese healers gained enduring and sometimes remarkably widespread followings.¹⁸⁸

The medical world saw major successes for Chinese doctors and their Caucasian patients. Li Putai,¹⁸⁹ Tan Fuyuan, Chang Yitang and Tan Feixuan (all of Los Angeles),¹⁹⁰ and Jin Yok Gong¹⁹¹ of San Francisco, and many Virginia City doctors (especially Hop Lock)¹⁹² developed favorable reputations among a white clientele that grew as exclusion laws diminished Chinese patients. Many of these doctors had large followings among white females with genealogical complaints.¹⁹³ Dr. Ah Sang gained respect and gratitude among white miners of the Sierra Nevada through his skillful healing abilities and hospital managerial skills.¹⁹⁴ The multi-generational Ah-Fong family of doctors enjoyed a much storied following in Idaho.¹⁹⁵ In

¹⁸⁴ Baker City, Oregon, BEDROCK DEMOCRAT, Feb. 7, 1872 at p.1, c.4 Canyon City, Oregon, BLUE MOUNTAIN EAGLE, Aug. 24, 1906 at 1, c.3-4 ; A.W. Loomis, *Medical Art in the Chinese Quarter*, 2:6 THE OVERLAND MONTHLY 496-506 (June 1869); William M. Tisdale, *Chinese Physicians in California*, 68 LIPPINCOTT'S MONTHLY MAGAZINE 411-416 (Mar. 1899).

¹⁸⁵ The *Blue Mountain Eagle* for 1909 alone is filled with many dozens of ads and fake news stories advertising unorthodox Western doctors. For specific examples, see Baker City, Oregon, BEDROCK DEMOCRAT, Jan. 31, 1872 at 4, c.6 ; Baker City, OREGON, BEDROCK DEMOCRAT, Apr. 17, 1872 at 4, c.5, c.6 ; Canyon City, Oregon, BLUE MOUNTAIN EAGLE, Sept. 30, 1910 at 2, c.1 ; Sarvis, *supra* note 113, at 66.

¹⁸⁶ Haiming Liu, *Resilience of Ethnic Culture: Chinese Herbalists in the American Medical Profession*, 1.2 J. ASIAN AM. STUDIES 180, 184-85 (1998); Tsai, *supra* note 23, at 38.

¹⁸⁷ For an example, see the sarcastic, racist account concerning Dr. Wing Lee's "failure" to cure a Mrs. L.A. Nye, who was in the final stages of tuberculosis in Canyon City in BLUE MOUNTAIN EAGLE Aug. 24, 1906, at 1, c. 3-4 .

¹⁸⁸ CHIA-LIN CHEN, A GOLD DREAM IN THE BLUE MOUNTAINS 124-25 (1972); Lee, *supra* note 117, at 204; MAGNAGHI, VIRGINIA CITY'S CHINESE COMMUNITY 128; Sarvis, *supra* note 113; STAN STEINER, FUSANG: THE CHINESE WHO BUILT AMERICA 123 (1980).

¹⁸⁹ Lu, *supra* note 183, at 178; (also spelled Li Po Tai, in Tisdale, *supra* note 184, at 412, 416).

¹⁹⁰ *Id.* at 178.

¹⁹¹ *Death of Legendary Ing Hay Closes Pioneer Story Era* in BLUE MOUNTAIN EAGLE, July 25, 1952 at 1, c.6 (noted for his reputation as perhaps the greatest Chinese doctor of his era). See also CHEN, *supra* note 188, at 122.

¹⁹² MAGNAGHI, *supra* note 188, at 128.

¹⁹³ JEFFERY BARLOW and CHRISTINE RICHARDSON, CHINA DOCTOR OF JOHN DAY 58 (1979); Chen, *supra* note 177, at 123; CHRISTOPHER MUENCH, ONE HUNDRED YEARS OF MEDICINE: THE AH-FONG PHYSICIANS OF IDAHO 65 (1984).

¹⁹⁴ STEINER, *supra* note 188, at 123.

¹⁹⁵ MUENCH, *supra* note 193, at 51-80.

January 1900, when the Idaho Supreme Court ruled in his favor,¹⁹⁶ C.K. Ah Fong became perhaps the only doctor practicing traditional Chinese medicine in the United States to gain a standard medical license. Beginning in the late nineteenth century in the small town of John Day (in remote eastern Oregon) Ing Hay became a legendary physician who practiced for nearly sixty years and eventually drew clients from as far away as San Francisco and Alaska.¹⁹⁷ Local white doctors, apparently jealous of Ing Hay's popularity, tried unsuccessfully on at least two occasions to have Ing Hay prosecuted.¹⁹⁸ Utilitarian friendship, as the name indicates, includes an inherent mutual self-interest. Although impossible to measure, sometimes these friendships clearly went beyond the conviviality associated with an immediate professional exchange, as some of the examples above clearly demonstrate.

IV. PROGRESSIVE ERA FRIENDSHIPS OF PLEASURE AND VIRTUE

Churches sometimes functioned as institutions providing an arena of benevolent exchange and friendships among Asians and Caucasians.¹⁹⁹ In 1853 Presbyterians Dr. and Mrs. William Speer began a San Francisco mission to the Chinese.²⁰⁰ In 1859 Mr. and Mrs. A.W. Loomis took charge, and began preaching the gospel and teaching English.²⁰¹ The Loomises illustrate a complex and sometimes ambiguous aspect of this study because of considerations of cultural imperialism²⁰² and Christian arrogance.²⁰³ The age of imperialism included a major Christian missionary component, in which (for the United States) mainly Protestants evangelized abroad.²⁰⁴ The idea of converting the heathen masses is almost as ancient as Christianity itself, going back to Paul the Apostle and his travels to spread what even-

¹⁹⁶ Ah Fong v. McCalla, 59 P. 930 (Idaho 1900).

¹⁹⁷ Barlow and Richardson's *China Doctor of John Day* is, unfortunately, heavily flawed. A better source is Chen's *A Gold Dream in the Blue Mountains*.

¹⁹⁸ Canyon City, Oregon, BLUE MOUNTAIN EAGLE, May 21, 1915 at 3, c.5; BLUE MOUNTAIN EAGLE, Nov. 23, 1928 at 1, c4 ; BLUE MOUNTAIN EAGLE, May 10, 1929, at 4, c.1; BLUE MOUNTAIN EAGLE, June 7, 1929; at 2, c.3; Sarvis, *supra* note 113, at 67.

¹⁹⁹ GUNTHER BARTH, BITTER STRENGTH 169 (1964); R. SEAGER, SOME DENOMINATIONAL REACTIONS TO CHINESE IMMIGRATION IN CALIFORNIA 53 (1959).

²⁰⁰ Seager, *supra* note 199, at 49.

²⁰¹ TSAI, *supra* note 23, at 43-44.

²⁰² Although he does not focus on the religious aspect, for the problematic definitions and usage of 'cultural imperialism', see JOHN TOMLINSON, CULTURAL IMPERIALISM: A CRITICAL INTRODUCTION 2-33 (1991).

²⁰³ The term 'cultural imperialism' has loose and variegated usages, but the disrespectful imposition of Christianity on non-Christian cultures cannot be denied, see generally KENNETH M. MACKENZIE, THE ROBE AND THE SWORD: THE METHODIST CHURCH AND THE RISE OF AMERICAN IMPERIALISM (1961); WILLIAM R. HUTCHISON, ERRAND TO THE WORLD: AMERICAN PROTESTANT THOUGHT AND FOREIGN MISSIONS (1987). For an inadvertent reflection of this biased orientation in the account of the Protestant minister William Speer, see generally SPEER, *supra* note 32, at 605-24.

²⁰⁴ *Id.* See also SEAGER, *supra* note 199, at 51.

tually became Christianity (Paul himself was not only a Jew, but a Pharisee).²⁰⁵ Unwittingly, in many cases, Christian missionaries thus became purveyors of cultural arrogance, for conversion to Christianity is absolute. This is because, according to the Hebrew Bible's classic requirement that has followed all monotheisms, the true believer must forsake "all other gods"²⁰⁶ before the One, whether that be Yahweh, God, or Allah.

Buddhism, on the other hand, historically spread far from its India birthplace in part because of its tolerance of established religions and thus a resulting polytheistic syncretism.²⁰⁷ For example, China's ancient tradition of Taoism heavily influence a school of Buddhism that came to be called *Ch'an* (*San* in Korean, but more familiar to Westerners in its Japanese pronunciation, Zen).²⁰⁸ By the time Japanese came to America they had lived for centuries as both Shintoists (the preceding tradition) and Buddhists (the imported foreign religion and philosophy).²⁰⁹ Buddhism had faded away as a popular religion in China, but Chinese continued to see no conflict in considering themselves both Confucianists and Taoists. A government official might see himself as a Confucianist by day and a Taoist by night, or a Confucianist during his career and a Taoist during retirement.²¹⁰ This difference in approach to religion helps explain why a foreign religion like Buddhism could be so successful, while Christianity was a comparative failure.²¹¹ The era of imperialism did not help matters. As Chinese scholar Monlin Chiang wrote, "Buddhism arrived in China on the back of an elephant. Christianity arrived on the deck of a gunboat."²¹² No

²⁰⁵ For a good introduction to Paul the Apostle, see GARRY WILLS, *WHAT PAUL MEANT* (2006).

²⁰⁶ *Exodus* 20:3. For an example of a historical application of this attitude, see LIAM M. BROCKEY, *JOURNEY TO THE EAST: THE JESUIT MISSION TO CHINA, 1579-1724*, at 60-61, 117, 296, 298-301 (2007).

²⁰⁷ ARTHUR F. WRIGHT, *BUDDHISM IN CHINESE HISTORY* 33, 36, 49, 62, 71, 97, 98-104 (1965).

²⁰⁸ *Id.* at 78-79; DAIGAN MATSUNAGA AND ALICIA MATSUNAGA, *FOUNDATION OF JAPANESE BUDDHISM* 194 (1974).

²⁰⁹ MATSUNAGA, *supra* note 208, at 55, 167-68, 174, 272, 294-95, 300-01.

²¹⁰ WILLIAM DE BARY, *SOURCES OF CHINESE TRADITION* 15, 48 (2d ed. 2000).

²¹¹ As recently as the 1950s Buddhism and Shintoism overwhelmingly represented religious adherents in Japan. For statistical tables, see WILLIAM K. BUNCE, *RELIGIONS IN JAPAN: BUDDHISM, SHINTO, CHRISTIANITY* 173-180 (2d ed. 1963). By 1991 still less than 1% of Japanese were Christians, see CAROLYN FRANCIS, *CHRISTIANS IN JAPAN* 134 (1991). Similar statistics characterize China, see ALAN HUNTER AND KIM-KWONG CHAN, *PROTESTANTISM IN CONTEMPORARY CHINA* 278 (1993); RICHARD MADSEN, *CATHOLIC CONFLICT AND COOPERATION IN THE PEOPLE'S REPUBLIC OF CHINA IN GOD AND CAESAR IN CHINA: POLICY IMPLICATIONS OF CHURCH-STATE TENSIONS* 93, 104 (Jason Kindopp & Carol Lee Hamrin eds., 2004).

²¹² The author first heard this quote translated by Dr. Young-tsu Wong in one of his Chinese history classes, ca. 1984. The quote comes from the book *Xi Chao, Jiang Mengling zhu* (*Tide from the West*) published by Monlin Chiang (1886-1964) in 1960. The author thanks his professor for remembering the source and refreshing his memory after all these years.

doubt Monlin Chiang was remembering Western imperialism²¹³ (particularly the gunboats of the British and French),²¹⁴ which included two opium wars (1839-1842)²¹⁵ and (1856-1860)²¹⁶ and the destruction of the Yuanming Yuan, one of the most magnificent gardens in world history.²¹⁷

Keeping all such considerations in mind, and sometimes in spite of such considerations, American churches sometimes became altruistic locales of interracial friendship among Asians and Caucasians. For example, Junro Kashitani, born in 1899 in southern Japan, got involved in various Christian social circles in the United States.²¹⁸ “I have made best friends in the world among genuine Christians here. I have two American good spiritual mothers who are better than my own parents, good many friends among young people who are dearer than my own kindreds [sic].”²¹⁹ Better than her own parents? This is an extremely strong statement that surely exceeded the bounds of mere politeness.

Mary Nobe was an American born Japanese who grew up among Caucasians in Los Angeles during the early twentieth century.²²⁰ She rarely associated with fellow Japanese and felt alienated from their social circles. Nobe found some of her closest white friends in her Presbyterian church’s Philathea class. “I have been with this Philathea class so much that I am one of them. I go to their homes, eat with them, sleep with them, and take part freely in all their activities.”²²¹ Someone as acculturated as Nobe, who associated almost only with whites, could only expect to find friendship in her chosen Caucasian social circles.

Louie Chin (Chin Ming Gum) was a Chinese labor contractor who worked as foreman for an Anacortes fish company from 1909-1915.²²² His death from kidney failure at the age of forty-eight in a San Francisco hospital made the front page of the local newspaper.²²³ Like all Chinese labor contractors, Chin acted as a liaison between cultures. At the very least, such a role necessitated a strong command of English. But Chin had gone further in his attempt to acculturate, for he and his family began attending Anacortes’s Westminster Presbyterian Church.²²⁴ In July 1915, the Westminster Presbyterian Church baptized the late Chin’s two sons and four

²¹³ IMMANUEL C.Y. HSÜ, *THE RISE OF MODERN CHINA* 92-105, 107-19, 139-150 (5th ed. 1995).

²¹⁴ *Id.* at 154-66, chs. 8-9, 215, 325-30.

²¹⁵ *Id.* at 184-93.

²¹⁶ *Id.* at 205-19 (also known as the “Arrow War”).

²¹⁷ *Id.* at 215, 308. For an in-depth study of the Yuanming Yuan’s history and destruction, see YOUNG-TSU WONG, *A PARADISE LOST: THE IMPERIAL GARDEN YUANMING YUAN* (2001).

²¹⁸ *Life History of Junro Kashitani*, SURVEY OF RACE RELATIONS, box 34, file 1, at 1, 3.

²¹⁹ *Id.* at 3.

²²⁰ *Mary Nobe Interview*, SURVEY OF RACE RELATIONS, box 25, file 81, at 1.

²²¹ *Id.*

²²² ANACORTES AMERICAN, 1 Apr. 29, 1915.

²²³ *Id.*

²²⁴ *Id.*

daughters.²²⁵ Apparently the widowed Mrs. Chin (Yin Look) mistook the baptism ceremony as one that received her and her children as formal members of the church.²²⁶ After discovering this mistake, the church congregation voted to accept the Chins as members.²²⁷ That the church and its members accepted the Chin family must reflect some degree of tolerance, if not complete acceptance.

Between 1923-1933, the Baptist Chung Mei Home for boys in Berkeley, California, provided assistance for some 200 Chinese orphans.²²⁸ In addition to receiving food, shelter, education, and clothing, the Chung Mei Home clearly indoctrinated the boys in Baptist ideology. Some of the boys grew up to become Baptist missionaries in China and the United States.²²⁹ But adult Asian interaction with American churches could be more complex, and charity was hardly the sole domain of mainstream Caucasian America. In fact, Japanese and Chinese had widespread reputations for helping white society with generous contributions of cash when others were in need.²³⁰

Public schools often became a major conduit of cultural and social contact between immigrant Asian families and mainstream American society.²³¹ Pre-war Washington and Oregon state law required school attendance of all children regardless of ethnicity.²³² In California there was much more school segregation.²³³ Japanese students were not allowed to attend “white” schools until after 1906, whereas the Chinese had to wait until the late 1920s.²³⁴ Where integrated, public schools — more than any other institution — became the location of Asian acculturation. Again, friendships did not necessarily develop. White children, especially if they had racist parents, often teased Asian children for being different.²³⁵ High academic

²²⁵ *Minutes of the Session and Church Register* (July 11, 1915) Westminster Presbyterian Church archives, Anacortes, Washington.

²²⁶ *Id.*

²²⁷ Westminster Presbyterian Church, Anacortes, Washington, THE NEW CHURCH REGISTER, Mar. 30, 1917.

²²⁸ Charles Shepard, *The Story of Lee, Wong and Ah Jing* (1885). This document appears to have been reproduced as Charles R. Shepherd, *The Story of Chung Mei* (1938).

²²⁹ *Id.*

²³⁰ SPEER, *supra* note 32, at 636; McComber, *A Japanese Agricultural Community*, *supra* note 148, at 4-5; A.R. Johnson *Interview*, *supra* note 140, at 3; Mr. Demarest *Interview*, SURVEY OF RACE RELATIONS, box 28, file 205, at 3.

²³¹ YAMATO ICHIHASHI, JAPANESE IN THE UNITED STATES 222 (1969); DAVID K. YOO, GROWING UP NISEI: RACE, GENERATION, AND CULTURE AMONG JAPANESE AMERICANS OF CALIFORNIA, 1924-49, at 29-30, 38 (2001).

²³² John E. Corbally, *Orientalism in the Seattle Schools*, 16:1 SOCIOLOGY & SOCIAL RESEARCH 61-67 (1931); ELIOT G. MEARS, RESIDENT ORIENTALS ON THE AMERICAN PACIFIC COAST: THEIR LEGAL AND ECONOMIC STATUS 352-54, 363 (1928).

²³³ DANIELS, ASIAN AMERICA, *supra* note 18, at 111.

²³⁴ JOYCE KUO, EXCLUDED, SEGREGATED, AND FORGOTTEN: A HISTORICAL VIEW OF THE DISCRIMINATION AGAINST CHINESE AMERICANS IN PUBLIC SCHOOLS IN CHINESE AMERICA: HISTORY AND PERSPECTIVES 39, 40 (2000).

²³⁵ ZHU, *supra* note 107, at 175.

achievements,²³⁶ reflecting a profound and enduring legacy of Confucian values of learning²³⁷ sometimes inspired jealousy among Caucasian children.²³⁸ White teachers and school administrators, on the other hand, not surprisingly almost uniformly praised their Asian students.²³⁹ Such teachers often befriended Asian parents who, in turn, sometimes bestowed gifts upon the teachers.²⁴⁰ Thus, school relationships extended beyond the school itself. Among the school children authentic white-Asian friendships formed, and sometimes endured for many years beyond high school.²⁴¹ For Nisei, in fact, public school sometimes provided a refuge from intergenerational family tensions. Tadao Kimura, born in 1907 in Seattle, reflected upon his public education shortly after graduating in 1924 as class valedictorian from Franklin High School:

My school life has been a far more pleasant one [compared to life outside of school]. In the grade school we formed very close friendships and had no trouble at all. In the high school I missed the close friendships but that was impossible because of the frequent change of classes. But the friendly feeling was there and I enjoyed high school life very much. As valedictorian I thought there might be considerable trouble. But I was mistaken. Except for the contempt of a few worthless pupils, everyone was sincere in his congratulations.²⁴²

No doubt in part because of their diligence, Asian students like Tadao Kimura often formed close attachments to their Caucasian teachers. “The

²³⁶ DANIELS, *ASIAN AMERICA*, *supra* note 18, at 173-74; SHOTARO FRANK MIYAMOTO, *SOCIAL SOLIDARITY AMONG JAPANESE IN SEATTLE* 54 (1984).

²³⁷ In the United States most scholars and general readers think of Confucius as a philosopher. On the other hand, traditionally the Chinese revered Confucius as a teacher. The annual Teacher’s Day holiday in Taiwan (where traditionalism escaped the purges of the mainland’s Cultural Revolution of the 1960s) has a magnitude somewhat equivalent to the American Thanksgiving holiday, and is only outranked in magnitude by the Lunar New Year celebration. Even the author, a *yang gue tze* (i.e., “foreign devil”) English teacher in Taipei, Taiwan (from 1987-1988) received gifts from his students on Teacher’s Day.

²³⁸ JEAN PAJUS, *THE REAL JAPANESE CALIFORNIA* 179, 182 (photo. reprint 1971) (1937); EILEEN SARASOHN, *THE ISSEI: PORTRAIT OF A PIONEER, AN ORAL HISTORY* 62 (1983).

²³⁹ KAZUO ITO, *supra* note 159, at 676-79; HERBERT JOHNSON, *supra* note 128, at 33-36; LOFTUS, *supra* note 128, at 83; MEARS, *supra* note 232, at 363; J.S. TOW, *THE REAL CHINESE IN AMERICA* 62-65, 72 (photo. reprint 1970) (1923).

²⁴⁰ LOWE, *supra* note 107, at 78-80; MIYAMOTO, *supra* note 236, at 53-54.

²⁴¹ Sui Sin Far, *Leaves from the Mental Portfolio of an Eurasian*, 66:3138 *THE INDEPENDENT* 125-32 (1909); *Ruby Hirose Interview*, *SURVEY OF RACE RELATIONS*, box 27, file 159, at 2; *Florence Kojima Interview*, *SURVEY OF RACE RELATIONS*, box 29, file 15, at 2; *Susie Yamamoto Interview*, *SURVEY OF RACE RELATIONS*, box 25, file 77, at 1; ZHU, *supra* note 107, at 175-76 (the latter source may provide a somewhat idealized account based exclusively on Caucasian recollections).

²⁴² Ruth H. Greiner, *The Japanese Valedictorian of Franklin High School*, in *SURVEY OF RACE RELATIONS* box 25, file 82 at 6-7. (A copy of the Tadao Kimura essay also appears in the *Smith Collection*, Series A, #163). For another Nisei account of intergenerational difficulties, see “G.S.” *Interview*, *SURVEY OF RACE RELATIONS*, box 25, file 79, at 2.

teachers took a great interest in me and my work as I was different,” remembered J. Lim, a Californian-born child of early twentieth century Korean immigrants.²⁴³ “Studying was not difficult for me for behind me lay generations of students and scholars,” Lim said, again reflecting the Confucian value of education that permeated all of East Asia. “In my school work I was aided more than handicapped because of my racial differences.”²⁴⁴

Marjorie McComber was an adamant friend to the Japanese in the Fullerton, California area.²⁴⁵ In 1924 she remembered a fourteen-year-old Japanese boy called George who attended their school. He was her first Japanese friend. McComber recalled, “He was made a brother to us all. He entered in all our games. We children never even had a thought that he should play by himself and live in a place in the world away from us. This racial conflict did not enter in our lives. He helped and respected us and we helped and respected him.”²⁴⁶

Hida Watanabe was born in Japan around 1907 and moved to the United States with her family about three years later.²⁴⁷ They lived in Texas, Colorado, and Missouri before moving to Los Angeles around 1921. “My friends were American girls and I had intimate friends,” she wrote in 1925. “I never felt different from them.”²⁴⁸ Unlike some of the other Asian students mentioned in this study, Watanabe did not participate in extracurricular school activities, so her friendships must have developed through general school attendance.²⁴⁹

Some of the early twenty-first century Caucasian elders of Anacortes, Washington, still remembered their former schoolmates Sumi and Fumi, daughters of Charles and Sakaye Tanikawa.²⁵⁰ Charles immigrated to the United States in 1907; Sakaye, twelve years later, the same year they were married.²⁵¹ Both worked as fish boners for the Matheson codfish plant. In 1940, their daughters were nineteen and sixteen years old. Sadly, the United States government interned the Tanikawa family at the particularly harsh Tule Lake camp during World War II. Heartbroken and betrayed, Charles Tanikawa moved the family to Japan after the war, where Fumi, the youngest daughter, soon died of a ruptured appendix. But the reason the elderly whites of Anacortes know and remember this tragic story is be-

²⁴³ Life History of J. Lim, SURVEY OF RACE RELATIONS, box 29, file 270, at 2.

²⁴⁴ *Id.* For other examples of the teacher-student bond, see the S. Yamadee Interview, SURVEY OF RACE RELATIONS, box 29, file 264, at 4; and an anonymous Chinese Interview, SURVEY OF RACE RELATIONS, box 29, file 268, at 3.

²⁴⁵ Marjorie McComber Interview, SURVEY OF RACE RELATIONS, box 26, file 117, at 2.

²⁴⁶ *Id.*

²⁴⁷ Hida Watanabe Interview, SURVEY OF RACE RELATIONS, box 31, file 314, at 1.

²⁴⁸ *Id.*

²⁴⁹ *Id.*

²⁵⁰ Wallie Funk Interview, (Oct. 16, 2001); Sarah Boynton Interview, (June 30, 2000) (both transcripts available in the Anacortes Museum); Virginia Boynton Perkins, telephone conversation with author (Dec. 11, 2001).

²⁵¹ 15th Census of the United States (1930).

cause they kept exchanging letters with the Tanikawas throughout the war and afterwards, as true friends would, of course.²⁵² Here is a clear example of “friendships of pleasure” that endured long after the circumstances that spawned them. This begins to suggest something closer to the nature of Aristotle’s “true friendship,” especially considering the detrimental factors of time and geographical distance, not to mention race-based internment.

Public school teachers and administrators sometimes found themselves at the center of racist controversies involving their Asian students. In the case of Tadao Kimura (described above), the *Seattle Star* took issue with a Japanese student being honored as valedictorian.²⁵³ In response, the school principal (a Mr. Reid) asked a teacher to write an editorial defending Kimura and the school in *Tolo*, the school paper.²⁵⁴

The *Seattle Star* raised another racist furor during the mid-1920s when Harrison School cast Fred Kosaka as George Washington in a school play.²⁵⁵ School principal Eugenie B. Parriseau defended the choice, citing popular student selection of Kosaka for the role. “I cannot see even now why anyone should object,” Parriseau protested; “being a citizen of this country why should he not play the role of the ‘Father of his country?’”²⁵⁶ The Sons of Veterans and the Women’s Auxiliary attacked Kosaka’s teacher, a Miss Waite (who ironically was a member of the Daughters of the American Revolution), and asked her to apologize or resign.²⁵⁷ Superintendent of Seattle Public Schools, a Mr. Willard, refused to ask the Board of Education intervene. Willard supported Waite and personally helped write a defensive response to her attackers.²⁵⁸

Good school relations continued into higher education, of course. Frank Ishi, a Long Beach High School graduate, enjoyed his college education at Stanford.²⁵⁹ “I mingled entirely with American boys and had an American room-mate,” he remembered in 1924. “I had a good time while there. I mingled freely in their social functions. I went to their dances where I was the only Japanese and danced with American girls.”²⁶⁰

²⁵² Funk *supra* note 250; Boynton *supra* note 250.

²⁵³ Greiner, *Japanese Valedictorian*, *supra* note 242, at 1-8.

²⁵⁴ *Id.* at 3.

²⁵⁵ Ruth H. Greiner, *Seattle’s Japanese George Washington*, in SURVEY OF RACE RELATIONS, box 27, file 164, at 2-3.

²⁵⁶ *Id.* at 3.

²⁵⁷ *Id.* at 4.

²⁵⁸ *Id.* at 5. Other cases were not so auspicious. In 1922 the Hollywood *Daily Citizen* helped agitate racism when the students at Le Comte Junior High School elected (with a 600 vote margin) John Aiso as student boy president. In this case, school administrators ultimately surrendered to public and media pressure and replaced the student form of government with a faculty council. The numerous *Daily Citizen* articles as well as interviews with participants are compiled in SURVEY OF RACE RELATIONS, box 29, file 266.

²⁵⁹ Frank Ishi Interview, SURVEY OF RACE RELATIONS, box 26, file 98, at 1.

²⁶⁰ *Id.*

While in Los Angeles, another anonymous Japanese graduate of Stanford carried on a lengthy correspondence between 1922 and 1923 with a certain Mrs. E. Snell of the Stanford placement office.²⁶¹ While technically trying to find the alumnus a position appropriate for his training in chemistry, Snell also apparently offered moral support and encouragement during a frustrating and disheartening time of lengthy unemployment or under-employment.²⁶² “Adelaide K.,” a Korean native, attended University of Southern California during the 1920s. “The friendships I have cultivated among the students on this campus for the last two years have been very precious and some of them are my real personal friends whom I talk to as tho [sic] they were my race,” she wrote.²⁶³

Laro Kanow attended University of Southern California’s engineering school and, through the civil service competition, became a draftsman for a city engineering office in the Los Angeles area.²⁶⁴ In a 1924 interview Kanow described himself as socializing mostly outside of Japanese circles, even though he was president of the Japanese Young Men’s Association in Long Beach.²⁶⁵ He criticized fellow Japanese for isolating themselves from Caucasians and maintained a personal attitude of determination in the face of adversity. “My relations in the engineering department of the city have been the very best,” he reported. “The other employees in the department treat me very well and I have encountered no unpleasantness whatsoever. I always go out to lunch with a group of the fellows.”²⁶⁶ But while Kanow favored social assimilation with the mainstream, he also wanted his behavior to reflect favorably on any future Japanese applicants²⁶⁷ — and in this regard very much reflected the group consideration of his Asian heritage.

Other friendships of pleasure involved Caucasians and Asians living in the same neighborhood.²⁶⁸ The fact that racial mixing could take place in residential neighborhoods at all was remarkable, considering the segregated South and the tendency toward ethnicity-specific urban immigrant neighborhoods in many cities. Non-racist whites lauded their Asian neigh-

²⁶¹ Anonymous Japanese student letters, SURVEY OF RACE RELATIONS, box 28, file 240.

²⁶² *Id.*

²⁶³ Adelaide K., *Autobiography*, in SURVEY OF RACE RELATIONS, box 30, file 33, at 2. Yung Wing, Lue Gim Gong, Huie Kin, No-Young Park, Emma Fong Kuno and her respective Chinese and Japanese husbands also reflected the generally greater amiability that contrasted college atmospheres from general American society. See YUNG WING, MY LIFE IN CHINA AND AMERICA 24-41 (1909). See also ROGER DANIELS, COMING TO AMERICA: A HISTORY OF IMMIGRATION AND ETHNICITY IN AMERICAN LIFE 247-48 (2002); KIN HUIE, REMINISCENCES chs. 7-8, (1932); NO-YONG PARK, CHINAMAN’S CHANCE 77-78, 85, 87, ch. 12 (3d ed. 1948); Emma Fong Kuno, *My Oriental Husbands*, SURVEY OF RACE RELATIONS, box 25, file 53.

²⁶⁴ Laro Kanow Interview, SURVEY OF RACE RELATIONS, box 29, file 269, at 1.

²⁶⁵ *Id.* at 3.

²⁶⁶ *Id.* at 2.

²⁶⁷ *Id.*

²⁶⁸ Liping Zhu describes some of the earliest examples pertaining to integrated mining camps of the mid-1860s. See ZHU, *supra* note 107, at 161, 167.

bors for their neatness, thrift, honesty and politeness — and sometimes emphasized their point by citing negative counter-examples among other Caucasians.²⁶⁹

Mrs. Ko Wing Kan grew up in a small town near Vancouver, British Columbia, where hers was the only Chinese family.²⁷⁰ She remembered many Caucasian women making friendly social calls on her mother. “I like the country people much better than the Canadians here [in the city],” she said in 1924 while being interviewed in Vancouver.²⁷¹ Susie Yamamoto had similar memories of San Bernardino after moving to Los Angeles.²⁷² In her San Bernardino high school she had been the only Japanese girl but was active in student affairs and remembered the other students treating her well. “Everybody in San Bernardino knew me and spoke to me. When I would go along the streets, the judge, district attorney, and mayor would speak to me. I was frequently called to court to act as interpreter.”²⁷³ Eileen S. Sarasohn remembered her Rainer Valley, Washington, neighbors as having been good to her family, to the point where the “neighbors competed in inviting my [Nisei] children to birthday parties.”²⁷⁴

An anonymous Fresno policeman (who had grown up in rural Vermont) countered the apparently prevalent criticism among California Caucasians of the 1920s about Asian women working in the fields.²⁷⁵ “I can remember in hay time and harvest, my mother and my sisters used to come into the field and rake hay and shock oats and work right along side of us men. It didn’t hurt them a bit.”²⁷⁶ The officer, whose beat was Fresno’s Chinatown, praised Asian industriousness. “These Japanese and Chinese young folks growing up will make good citizens,” he said.²⁷⁷

During the 1920s and 1930s, when demographics began to reflect the 1882 Chinese Exclusion Act, small town newspapers in the Pacific Northwest turned wistful at the disappearance of their elder Chinese. “The Last

²⁶⁹ McComber Interview, *supra* note 245, at 4; McComber, *A Japanese Agricultural Community*, *supra* note 148, at 1; Masa Higashida Interview, SURVEY OF RACE RELATIONS, box 27, file 158, at 4; St. Paul and Tacoma Lumber Company general information, SURVEY OF RACE RELATIONS, box 28, file 202, at 2; Mr. and Mrs. Ben Dick Interview, SURVEY OF RACE RELATIONS, box 28, file 203, at 3, 5; Olson, *Oriental in the Lumber Industry*, *supra* note 151, at 6; Mr. Gehrkins Interview, SURVEY OF RACE RELATIONS, box 29, file 254, at 3, 7-9; Mr. and Mrs. Sessions Interview, SURVEY OF RACE RELATIONS, box 29, file 5, at 2-3; SARASOHN, *supra* note 238, at 62; Chloe Holt, *Interview with Spaulding, Hollywood*, in, SURVEY OF RACE RELATIONS, box 29, file 258, at 1-2.

²⁷⁰ Ko Wing Kan Interview, SURVEY OF RACE RELATIONS, box 24, file 13, at 4.

²⁷¹ *Id.* An anonymous Chinese woman gave a similar account in *Story of a Chinese Girl Student*, in SURVEY OF RACE RELATIONS, box 25, file 56, at 3).

²⁷² Susie Yamamoto Interview, SURVEY OF RACE RELATIONS, box 25, file 77, at 1.

²⁷³ *Id.*

²⁷⁴ SARASOHN, *supra* note 238, at 62.

²⁷⁵ *Chinatown Police Squad, Fresno, California*, in, SURVEY OF RACE RELATIONS, box 28, file 228, at 1-2.

²⁷⁶ *Id.*

²⁷⁷ *Id.*

of His Type,” reported eastern Oregon’s *Blue Mountain Eagle* in 1922 when “China Gyp” died in Canyon City.²⁷⁸ The front page story painted a sympathetic portrait of an individual isolated by language and cultural barriers and (in parlance common to such occasions) described him as “quite a character.”²⁷⁹ But then the writer continued, praising “Gyp” as “a good citizen, peaceful, law abiding, friendly and industrious.”²⁸⁰ Seven years later the final Chinese resident of Canyon City, an octogenarian called “China How,” departed for Asia with his “old time friend” Dr. J.H. Fell escorting him to Seattle.²⁸¹ Despite the reporter’s racist overtones, he nevertheless admitted that Canyon City residents themselves thought well of How, regretted his departure, and took up a collection to help pay for his ocean passage.²⁸² They knew they would never see him again. Later in 1930 the *Anacortes American* sung the virtues of Chin Toy upon his death.²⁸³ “Like most of his race, he was an honest man whose promise was equivalent to the mortgage bond of the general run of up-to-date white men. He paid his way to the end,”²⁸⁴ the paper reported in a remarkably disparaging comparison to the dominant race. Finally, in 1939, Grant County, Oregon’s “China Sam” died at 82. The newspaper writer remembered him as a patient and industrious worker, and again, isolated in the dominantly white society.²⁸⁵ All of these stories strike a regrettable note today, for they seem to reflect what were then lost opportunities. But, of course, they could have omitted writing these stories in the first place. Perhaps the aging Chinese, in this case, had won some well-deserved respect, even if belatedly.

Competitive adult sports also encouraged respect among whites and Asians.²⁸⁶ Asian boxers, tennis players, bicycle racers, golfers, baseball, football, and basketball teams all competed against Caucasians. Often they won respect and even affection. The sports press was no exception in racist depictions, and yet felt compelled to praise Ah Sing, a prize fighter of the early 1900s.²⁸⁷ George Yamauchi and Harry Honda, former baseball players, made the typical observation that athletic competition tended to dissolve discrimination, momentarily, if not longer.²⁸⁸ The white community of Wapato, Washington, even adopted the local Japanese American baseball team as their home team, which became the pride of both Japanese

²⁷⁸ BLUE MOUNTAIN EAGLE, Nov. 24, 1922, at 1, c.1.

²⁷⁹ *Id.*

²⁸⁰ *Id.*

²⁸¹ BLUE MOUNTAIN EAGLE, Dec. 27, 1929, at 1, c.6.

²⁸² BLUE MOUNTAIN EAGLE, Jan. 3, 1930, at 1, c.6.

²⁸³ ANACORTES AMERICAN, June 26, 1930, at 6.

²⁸⁴ *Id.*

²⁸⁵ BLUE MOUNTAIN EAGLE, Apr. 21, 1939, at 1, c.4.

²⁸⁶ DEENA K. NAKATA, *THE GIFT: THE OREGON NIKKEI STORY RETOLD* 49-50 (1995).

²⁸⁷ Joel S. Franks, *Chinese Americans and American Sports, 1880-1940*, in *CHINESE AMERICA: HISTORY AND PERSPECTIVE*, 135 (1996).

²⁸⁸ Michael L. Mullan, *Ethnicity and Sport: The Wapato Nippons and Pre-World War II Japanese American Baseball*, 26 *J. SPORT HIST.*, 106-7 (1999).

and Caucasians, especially during the 1934 and 1935 winning seasons.²⁸⁹ The Japanese lumber workers at Eatonville, Washington, formed a baseball team that encouraged inter-racial friendliness among players and spectators alike.²⁹⁰ J. Lim, the Korean high school student mentioned above, lettered in football, baseball, basketball and tennis — and even made captain of some teams. For his athletic and academic abilities Lim felt his classmates and friends respected him.²⁹¹ Kunitaro Yamada, Broadway High School student in Seattle summed up the potential effects of sports and other associations of pleasure:

As to the forms of associations that bring the most harmonious adjustment is the education in same school, various sports, association in church and in musical performances allowing both Americas and Japanese to hear or participate. In these places, we forget temporarily any feelings of race consciousness, and often arouse enthusiasm to go on to the high ideal in harmony.²⁹²

No doubt some of the relationships stemming from sports, school, and church blossomed into true friendship. Then, as now, personal “friendships of virtue” were the rarest of Aristotle’s three categories.²⁹³

During the early 1920s H. Fukasu was a college student in southern California, where he lodged with a Caucasian family.²⁹⁴ Fukasu had no intentions of befriending white people, but unexpectedly developed a close friendship with all three members of this family; the father, mother, and their 25-year-old son. After six months they became “real friends,” as Fukasu described it.²⁹⁵ “I forgot to think that they are Americans. Our friendship has continued until present and I call them American father and mother and they call me my child.”²⁹⁶ This friendship with individuals was all the more remarkable considering Fukasu’s detailed criticism of the United States in general. Fukasu denounced America for pseudo-Christian hypocrisy, imperialism, militarism, greed, international irresponsibility abroad and anti-Asian racism at home.²⁹⁷ But obviously he was not a rac-

²⁸⁹ Gail M. Nomura, *Beyond the Playing Field: The Significance of Pre-World War II Japanese American Baseball in the Yakima Valley*, in, BEARING DREAMS, SHAPING VISIONS: ASIAN PACIFIC AMERICAN PERSPECTIVES, 15-31 (1992).

²⁹⁰ OLSON, *Impressions, Attitudes and Incidents*, *supra* note 150, at 4-5, 22. See also Joseph R. Svinth, *Carving Out A Place: Japanese Americans in Eatonville, 1904-1942*, 16 COLUMBIA, 24-28 (2002).

²⁹¹ *Life History of J. Lim*, *supra* note 243, at 2.

²⁹² Kunitaro Yamada to Ruth H. Greiner, SURVEY OF RACE RELATIONS, box 27, file 151, at 3 (1924).

²⁹³ NICOMACHEAN ETHICS, *supra* note 139, at 212.

²⁹⁴ H. Fukasu, *Life History of an Oriental Student*, in, SURVEY OF RACE RELATIONS, box 32, file 14, at 5 (1924).

²⁹⁵ *Id.*

²⁹⁶ *Id.*

²⁹⁷ *Id.* at 7-9.

ist, and distinguished governmental and social policies from personal friends.

That Fukasu's close friendship would develop in a familial setting wherein he developed an "American father and mother" should not be surprising considering Confucius's strong emphasis on filial piety²⁹⁸ and harmonious family relations as the key to all other social relations.²⁹⁹ Millennia before contemporary American political rhetoric championed "family values" Confucius had already made insightful observations about the crucial importance of humanity's smallest and most fundamental social unit.³⁰⁰ Fukasu was hardly the only one to experience friendship in these circumstances.

Many people would argue that any happy marriage must be based on the deepest of personal friendships. No doubt a loss of friendship precedes many divorces. Of course, at the most intimate level marriage remains a "sacred mystery"³⁰¹ unknowable to those outside the marriage, and perhaps somewhat mysterious even to the spouses. In any case, obviously the "true nature" of this intimate relationship presents a challenge to the historian attempting to document interracial friendship. On the other hand, friendship does not lend itself to scientific measurement in the first place. Still, extrapolating that at least some of the interracial marriages then (as now) were based in true love seems reasonable, particularly when repeated and widespread firsthand testimony supports such a conclusion. Here are some examples.

Grace Shelp Horikoshi of Hollywood, California, grew up in Iowa and met her future husband while teaching English to Japanese immigrants.³⁰² They knew each other for seven years before they married, and Mrs. Horikoshi felt divine providence had blessed their union. "God made of one blood all nations," she said in a 1924 interview.³⁰³ By then she was learning Japanese. "I have never regretted our marriage. We have now been married five years and it grows better all the time."³⁰⁴

Huie Kin, a New York City pastor, married Louise Van Arman (daughter of a Troy, New York industrialist) and raised nine children.³⁰⁵ Mrs. Yip Quong, the first white woman married to a Chinese man in Van-

²⁹⁸ ANALECTS OF CONFUCIUS §§ I:6, II:5-8, II:20, IV:18, IV:25, XIII:18 (Waley, trans., Macmillan 1938).

²⁹⁹ *Id.* at 86-87, 114.

³⁰⁰ Confucius observed that the family (as the smallest social unit) must be orderly if a society at large were to enjoy order. See CONFUCIAN ANALECTS, THE GREAT LEARNING & THE DOCTRINE OF THE MEAN §§ 4,5, (Legge, trans. & ed. 2013).. This teaching was carried on by Mencius (372-289 B.C.), probably the most famous and influential Confucian philosopher following Confucius himself. See THE GREAT LEARNING AND THE MEAN-IN-ACTION 179-80 (E.R. Hughes, trans., E.P. Dutton & Co., 1943).

³⁰¹ *Ephesians* 5:32.

³⁰² Grace Shelp Horikoshi Interview, SURVEY OF RACE RELATIONS, box 28, file 235, at 3.

³⁰³ *Id.*

³⁰⁴ *Id.* at 5.

³⁰⁵ HUIE, *supra* note 263, at ch. 6.

couver (where she arrived in 1904), got on well with both Chinese and Caucasians.³⁰⁶ Max Sui Haw, Seattle resident of the 1920s, had no objection to marrying a white woman as long as she was as educated as himself.³⁰⁷ R. Kado, owner of the Sunnyland Nursery in Los Angeles during the 1920s, said that both white and Japanese communities accepted the marriage between a local Japanese man and his Caucasian wife.³⁰⁸ There were innumerable other examples of Asian-Caucasian interracial marriages or those favoring these unions.³⁰⁹ In an ironic twist of the pseudo-science of the day, some even speculated that children of such “mixed” unions were more intelligent than “pure” children of either race.³¹⁰

All marriages have varying degrees of happiness and harmony and sometimes grow stronger or weaker. Pre-war interracial marriages were probably subjected to more overt racism than in recent years. But to assume that none of these marriages involved true friendship is as absurd as assuming that they were all perfectly happy unions.

During the early 1890s Emma Ellen Howse and Walter Ngong Fong both attended Stanford University. According to Howse, Fong was the only Chinese student at Stanford, and quite popular with many other students.³¹¹ Howse and Fong, however, shared a deeper attraction and decided to marry. In 1897 they circumvented California’s anti-miscegenation laws by getting married in Denver, Colorado. They returned and made Berkeley their home.³¹² Walter Fong went into law practice and then taught

³⁰⁶ Yip Quong Interview, *Smith Collection*, Series A, #11.

³⁰⁷ Max Sui Haw Interview, *Smith Collection*, Series A, #244.

³⁰⁸ R. Kado Interview, *Smith Collection*, Series A, #302.

³⁰⁹ D. A. Smith Interview, *Smith Collection*, Series A, #4; SAN FRANCISCO CHRONICLE, June 20, 1897, at p.20, c.3; SAN FRANCISCO CALL, Dec. 21, 1921; SAN FRANCISCO EXAMINER, Mar. 20, 1923; Kuno, *My Oriental Husbands*, *supra* note 263, at 1, 3-4; Hideo Oyama survey response, SURVEY OF RACE RELATIONS, box 25, file 60, at 7; Mr. and Mrs. Toyotomi Ujimasa, *Case Brief on Anglo-Japanese Marriage*, in SURVEY OF RACE RELATIONS, box 25, file 62; box 26, file 104; box 26 file 121, at 5; Mrs. S. Sasabe Interview, SURVEY OF RACE RELATIONS, box 27, file 174; K. Fujita Interview and Mr. Okabe Interview, SURVEY OF RACE RELATIONS, box 28, file, at 7, 13; J.K. Fukushima Interview, SURVEY OF RACE RELATIONS, box 28, file 227, at 2; Dr. F.T. Nakaya Interview, SURVEY OF RACE RELATIONS, box 28, file 247, at 12; Anonymous Chinese Interview, SURVEY OF RACE RELATIONS, box 29, file 18, at 5. *See also* Daniel Liestman, *Utah’s Chinatowns: The Development and Decline of Extinct Ethnic Enclaves*, 64 UTAH HIST. Q. 271 (1996). Perhaps the most celebrated Asian-Caucasian marriage of the nineteenth century centered around Polly Bemis; for a newspaper story documenting her deathbed testimony *see* the PORTLAND OREGONIAN, Nov. 5, 1933, § 5 at 1, c. 8. For examples of a more ambiguous attitudes toward interracial marriage *see* ANACORTES AMERICAN, Feb. 6, 1902, at 2; CHAN, THIS BITTERSWEET SOIL, *supra* note 110, at 196; YOO, *supra* note 231, at 78, 83-86.

³¹⁰ SAN FRANCISCO EXAMINER, Nov. 11, 1922; Mar Sui Haw Interview, SURVEY OF RACE RELATIONS, box 28, file 244, at 3-5, 7; Kuno, *My Oriental Husbands*, *supra* note 263, at 25.

³¹¹ Kuno, *My Oriental Husbands*, *supra* note 263, at 2.

³¹² *Id.* at 4, 5. *See also* SAN FRANCISCO CHRONICLE, June 20, 1897, at 20, c.3.

Chinese at University of California. The Fongs later traveled to China where Walter tragically died of bubonic plague, leaving his wife with two sons to raise.³¹³ Emma returned to the United States and eventually married her late husband's friend and colleague, Yoshi S. Kuno, who had taught Japanese at University of California.³¹⁴ Around 1921 she wrote, "I became the wife of Professor Kuno, a strong bond between us being affection for the one that was gone and a desire to make a home for the children."³¹⁵

Emma Fong Kuno was understandably closer to her first husband but enjoyed a happy marriage with her second husband as well. "It is a lack of character that makes marriage a failure regardless of race," she wrote. "Other things can be borne with or overcome."³¹⁶ She attributed a happy social life among university community Caucasians, Chinese, and Japanese in Berkeley to her marriages and became the informal "mother" to various Chinese and Japanese student groups.³¹⁷

Beyond the most intimate of friendships we should not forget the innumerable incidental kindnesses (what Aristotle might have called simple "good will") shown between Asians and Caucasians. According to Aristotle, good will was or could be an initial stage toward friendship; certainly friendship could not develop without it.³¹⁸ Moreover, he wrote, "generally good will occurs because of excellence, or a kind of decency, where one person appears to another a fine character, or courageous, or something like that."³¹⁹ While almost always fleeting, the pleasure and happiness these instances of good will brought, even if momentary, was possibly equal in magnitude to incidental racism and hatred which caused commensurate pain and suffering. Early twentieth century examples include Hideo Tashima being included in evening board games with his Seattle landlord's family.³²⁰ White street car riders in Los Angeles came to a Japanese dentist's defense when the latter was attacked because of his race. A certain Mrs. Carrier's Seattle University District rooming house became a preferred enclave of Japanese girls.³²¹ An elderly Vancouver lady gave the teenaged Ko Wing Kan English lessons late into the night.³²² Chinese throughout the American West had a reputation for being kind to white children, and often bestowed gifts of candy upon them.³²³ Chinese and

³¹³ Kuno, *My Oriental Husbands*, *supra* note 263, at 6.

³¹⁴ *Id.* at 7.

³¹⁵ *Id.*

³¹⁶ *Id.* at 26.

³¹⁷ *Id.* at 26-31.

³¹⁸ NICOMACHEAN ETHICS, *supra* note 139, at 231, 232.

³¹⁹ *Id.* at 232.

³²⁰ Hideo Robert Tashima Interview, SURVEY OF RACE RELATIONS, box 24, file 38, at 3-4.

³²¹ Chiye Shigemura Interview, SURVEY OF RACE RELATIONS, box 24, file 44, at 1.

³²² Kan, *supra* note 270, at 2.

³²³ ANACORTES AMERICAN, July 28, 1898, at 3; CHAN, THIS BITTERSWEET SOIL, *supra* note 110 at 362-63; Edith Robinson Interview, Mar. 8, 1973, at 7-9; Polly M. Bell, *A Pioneer Woman's Reminiscences of Christmas in the Eighties*, 49 OR. HIST. Q. 284, 284-96 (1948); THERESA TREBON, CULTURAL ARTIFACTS OF EBEBY'S LANDING NATIONAL

Caucasian children alike regularly gathered around King Sing, an elderly Chinese doctor of early twentieth century Seattle who was apparently very kind hearted. When a certain white man died in the neighborhood, King Sing supported the widow and her children until the latter were old enough to begin earning some money of their own.³²⁴

Mrs. Peter Mayberg lived in the Asian section of Seattle for many years during the early twentieth century.³²⁵ She remembered her Chinese neighbors surnamed Sin, who tragically lost a three-year-old son when he was accidentally scalded to death. Mayberg was sensitive to the Chinese aversion to handling corpses, so she laid out the little boy in silk. From then on Mr. Sin would not accept payment from Mayberg when she shopped at his store. Finally, out of embarrassment, Mayberg had to stop shopping there altogether, for she felt the gifts were too generous.³²⁶

In addition to incidental kindness, a great deal might be said about the attitude of individuals involved in interracial friendships or potential friendships. Attitude, as mentioned, is a key component to moral cosmopolitanism.³²⁷ Kyo Inouye, a Los Angeles teenager in 1925, fondly remembered her Caucasian friends from high school.³²⁸ While she only counted one of them as an intimate friend, she enjoyed good relations with many others in the glee club and as captain of the baseball team.³²⁹ “You have to be friendly yourself, then they will be friendly too,” she wrote, and attributed certain instances of interracial indifference to Japanese reserve.³³⁰ Mrs. Florence Kojima, General Secretary of the Japanese YWCA in Los Angeles, said in 1924, “We of the present generation are the pioneers and I think the pioneers always have to suffer the hardships. It is for us to hope and work for better feeling. We know and keep telling our people that this [racism] is not the true spirit of America, that the true spirit of America is kind and fine and friendly.”³³¹ She added that eternal observation of non-racists everywhere, “When people come to know each other, they learn that people are all the same, the color of the skin makes no difference.”³³²

White friends of Asians objected to discrimination on the widest grounds of American nationalism. In 1920, Marjorie McComber (friend of the Japanese mentioned above) evoked the most idealistic notions of America to protest anti-Asian racism. “America is supposed to be the world’s

HISTORICAL RESERVE 482, 493, 524 (1998) (cultural resource management document deposited with the Wing Luke Asian Museum, Seattle, Washington); WHIDBEY NEWS-TIMES, Aug. 18, 1999, at A1.

³²⁴ Mrs. Peter Mayberg Interview, SURVEY OF RACE RELATIONS box 27, file 154, at 2.

³²⁵ *Id.* at 1.

³²⁶ *Id.*

³²⁷ APPIAH, *supra* note 5 at xiii, xv; Vertovec & Cohen, *supra* note 5, at 13; FORD, RACIAL CULTURE, *supra* note 5, at 164.

³²⁸ Kyo Inouye Interview, SURVEY OF RACE RELATIONS, box 31, file 6, at 3.

³²⁹ *Id.*

³³⁰ *Id.*

³³¹ Florence Kojima Interview, SURVEY OF RACE RELATIONS, box 29, file 15, at 3.

³³² *Id.*

greatest democratic nation ... but does she really show it with the present problem of her true neighbor called Japan?"³³³ McComber advocated the classic Jeffersonian idea of "Democracy through Education."³³⁴ And again, certain people in the legal world objected to discriminatory laws on moral grounds. In the 1920s, when American nativism had seen its nadir, a certain Thomas Wilson of the San Francisco area said, "No legislation has ever been attempted in this country that pointed out a logical reason why Chinese should be discriminated against more than Englishman, Germans or people of any other European nation."³³⁵ Seattle attorney Henry A. Monroe agreed. "I think the new immigration law is a most unfortunate one," he said. "The friendship of the Chinese should be cultivated for the good of mankind and certainly for the advantage of the United States commercially."³³⁶

Pre-World War II benevolence among Asians and Caucasians carried on through the war and afterward, of course. So much attention has focused on Japanese-American internment and subsequent, belated redress that many have forgotten that a minority of Caucasians defended their fellow Americans of Japanese descent amidst the actual relocation.³³⁷ For example, some former missionaries to Asia and West Coast university people spoke out against the "evacuation." Others were merely acquaintances. In early June 1942, approximately 163 people of Japanese descent from a number of northwestern Washington counties gathered in Mount Vernon, Washington, for "evacuation" to internment camps. "A large majority of the Japanese being evacuated were young persons born, raised and educated in local communities," the *Anacortes American* reported sympathetically (remarkably, given its racist reporting in earlier decades). "Large numbers of their friends and neighbors were on hand at the station to see them off."³³⁸ And while the overwhelming majority of Japanese-Americans were forced to sell their property at scandalously low prices, or lose such property through foreclosure, a few friendships involved whites who served as temporary caretakers for Japanese-American property throughout the war.³³⁹

³³³ McComber, *A Japanese Agricultural Community*, *supra* note 148, at 6.

³³⁴ McComber Interview, *supra* note 245, at 6.

³³⁵ Anonymous Chinese author, *The History and Problem of Angel Island*, in SURVEY OF RACE RELATIONS, box 27, file 150, at 6.

³³⁶ Henry A. Monroe Interview, SURVEY OF RACE RELATIONS, box 28, file 219, at 4.

³³⁷ Robert Shaffer, *Opposition to Internment: Defending Japanese American Rights during World War II*, 61 THE HISTORIAN 3, 597-619 (Spring 1999). See also BEYOND BARBED WIRE (Mac & Ava Films 1997).

³³⁸ ANACORTES AMERICAN, June 4, 1942, at 4.

³³⁹ Marta and Karl Olsen, helpless to aid the Sones of Seattle, nevertheless remained loyal friends. See SONE, *supra* note 147, at 164.

V. THE IMPORTANCE OF THE MORAL COSMOPOLITAN PRECEDENT

For all the damage that historic bigots accomplished, the unpopular few who went against their contemporary racism and nativism set important socio-legal precedents for what would eventually follow: a much richer social and economic integration of Asians and Asian-Americans into mainstream United States society, as well as less discriminatory immigration policies. As many scholars have pointed out, “precedent” as normally understood in law actually underlies and supersedes formal legalism, at least in a system of democratic constitutionalism, such as that of the United States. Thus, the popular will (including dissents) easily shapes constitutional law, and thereby also the nation’s character, even if indirectly and over protracted periods of time.³⁴⁰ One scholar even declared that “nonjudicial precedents are even more important than judicial precedents in shaping national identity.”³⁴¹ This line of activity and theory would generally fall into the tradition of what is variously called “populist” or “customary” constitutionalism.³⁴² In this context, judicial review becomes part of a step-by-step process whereby popular will, legislative actions, and the Supreme Court all interact, often over long periods of time, to realize the popular will’s long view (as opposed to momentary prejudices) which, in the best scenarios, comes to confirm Enlightenment principles of universal human rights.³⁴³ Populist constitutionalism is where moral cosmopolitans and nationalism have met and overlapped.³⁴⁴ Where populist constitutionalism got it wrong, as in the *Korematsu* decision and its widespread support,³⁴⁵ the moral cosmopolitan dissent became all the more important.³⁴⁶

³⁴⁰ RICHARD THOMPSON FORD, RIGHTS GONE WRONG, 127-29, 158 (2012); Anthony Lewis, *Historic Change in the Supreme Court*, N.Y. TIMES, June 17, 1962 at 7, 36, 38-39; SANDRA DAY O’CONNOR, THE MAJESTY OF THE LAW: REFLECTIONS OF A SUPREME COURT JUSTICE 166 (2004); WILLIAM G. ROSS, FORGING NEW FREEDOMS 206 (1994); MARK TUSHNET, I DISSENT: GREAT OPPOSING OPINIONS IN LANDMARK SUPREME COURT CASES xx, xxiv-xxv, 222, 223 (2008); MICHAEL J. GERHARDT, POWER OF PRECEDENT 3, 6, 171, 173, 175 (2008); MARK TUSHNET, TALKING THE CONSTITUTION AWAY FROM THE COURTS 182, 222n.12 (1999); Mark Tushnet, *Populist Constitutionalism as Political Law*, 81 CHI-KENT L. REV. 991, 996-97, 999-1000 (2006); BARRY FRIEDMAN, WILL OF THE PEOPLE 4, 177, 196, 205, 236, 367, 368, 381-384 (2010); Jeremy Webber, *A Judicial Ethic for a Pluralistic Age*, in MULTICULTURALISM AND LAW 88, 93, 98 (2007).

³⁴¹ GERHARDT, *supra* note 340, at 171.

³⁴² This is a common theme of *I Dissent, People Themselves, Talking the Constitution Away from the Courts*, and *Will of the People*.

³⁴³ LARRY D. KRAMER, PEOPLE THEMSELVES: POPULAR CONSTITUTIONALISM AND JUDICIAL REVIEW 208, 209 (2004); FRIEDMAN, *supra* note 340, at 238, 385.

³⁴⁴ TUSHNET, *supra* note 340, at 181, 191.

³⁴⁵ FRIEDMAN, *supra* note 340, at 372-73, 382.

³⁴⁶ TUSHNET, *supra* note at 340, at xix-xx, 124; KRAMER, *supra* note 343, at 332. GERHARDT, *supra* note 340, at 171; TUSHNET, *supra* note 340, at 191, 194.

The interracial friendships and examples of benevolence illustrated in this piece form sort of a backdrop of largely unorganized social dissent from the era's prevailing racism and nativism. There were, however, many examples of organized and institutionalized dissent that also reflected the moral cosmopolitan precedent, some of which appeared in various Supreme Court cases. Between Congressional institutionalization of Chinese exclusion in 1882 and the World War II era, Supreme Court cases dealing with Asians and Asian Americans involved important constitutional principles, yet in the broad view seem almost like legal technicalities within a context of overall societal acceptance of racial discrimination. And yet these dissents, no matter how momentarily inconsequential they may have seemed at the time, nevertheless offered continuity from Enlightenment ideals of universal rights and their eventual, more full realization during the late twentieth century.

In the much-noted case of *Yick Wo v. Hopkins* (1886) the U.S. Supreme Court declared a San Francisco ordinance in violation of the 14th Amendment because it specifically targeted Chinese laundry workers.³⁴⁷ The ordinance flagrantly discriminated against Chinese laundry operations without explicitly saying so, by declaring all wooden structures housing laundry facilities illegal.³⁴⁸ Few or none of the 80-plus Chinese laundries operated in masonry buildings, and yet some had been in operation for more than twenty years without incident. The city had reported no violation of fire regulations upon inspecting the many wooden laundries.³⁴⁹ The Supreme Court recognized the city ordinance for what it was, a would-be building code stipulation that actually unconstitutionally targeted a specific demographic. *Yick Wo v. Hopkins* was an unambiguous ruling against the prevailing racism of the era, and thus a ray of hope.

In *U.S. v. Jung Ah Lung* (1888), the defendant gained re-entry into the United States despite having lost his identification certificate.³⁵⁰ Somewhat remarkably, a California district court had originally granted Jung Ah Lung a writ of habeas corpus and ordered his release. Custom authorities balked and appealed the case to a California circuit court, which affirmed the district court's decision. The U.S. then appealed to the Supreme Court, where it lost, despite a dissent focusing on technicalities of the 1882 Exclusion.³⁵¹

In 1891, the Supreme Court denied sixteen-year-old Quock Ting entry into the United States. Custom officials detained Quock Ting in San Francisco and cited the 1882 Exclusion Act as grounds for denying admittance. Quock Ting and his father testified that he had been born in San Francisco, traveled to China at age ten, then tried to return to the United States six years later. As mentioned earlier, the Court decided Quock Ting's and his

³⁴⁷ *Yick Wo v. Hopkins*, 6 S.Ct. 1064 (1886).

³⁴⁸ *Id.* at 1066.

³⁴⁹ *Id.* at 1066-67.

³⁵⁰ *United States v. Jung Ah Lung*, 124 U.S. 624, 636 (1888).

³⁵¹ *Id.* at 623-24, 636-40.

father's sworn testimony were insufficient for establishing what would have been his birthright citizenship.³⁵² From Justice David Josiah Brewer's perspective, as expressed in his dissent, the Court's majority opinion failed to recognize the testimony as valid specifically because the witnesses were Chinese.³⁵³ In other words, Brewer's dissent echoed the rectification of the then-recent past of barring Chinese testimony in California state courts. A later generation would have called this "racial profiling," and certainly the dissent would prevail. But Brewer was only initiating his legacy of moral cosmopolitan dissent that would reverberate for many decades to come. Brewer had an unusual perspective for his time, and appreciated what we would now call anthropological relativism. In *Fong Yue Ting* (1893) Brewer reflected upon the legacy of legal discrimination of "the foremost Christian nation" against the Chinese, then wondered rhetorically about the irony of Christian missionaries seeking converts in China.³⁵⁴ It would be Brewer's dissent in *Fong Yue Ting* that would prove to be his greatest legacy regarding immigration law in general and deportation in particular.

Fong Yue Ting v. U.S. (1893) involved the expulsion of three Chinese men who had apparently failed to obtain certificates of residence, as required by the 1892 Chinese Deportation Act.³⁵⁵ The Act, incidentally, required at least one white witness to attest for the residency of an alien — and so federal law perpetuated California's forty-year old precedent, despite later amendment.³⁵⁶ In any case, Justice Horace Gray wrote the Court's majority opinion and described the various ways the three petitioners had failed to obtain certificates of residence. He affirmed a lower court's dismissal of writs of habeas corpus, stating that such action was consistent with both international law and United States constitutional law.³⁵⁷ On one level, this conclusion may have seemed straightforward, even purely procedural. But Justice David Brewer saw it in a different light. First, he dissented from the majority because he saw the Chinese men as legal residents and therefore constitutionally protected. More importantly, he saw the 1892 Chinese Deportation Act as denying due process, and therefore unconstitutional.³⁵⁸ To Brewer, the 100,000 Chinese living in the United States at that time were "not travelers, but resident aliens."³⁵⁹ Then, Brewer went on in eloquent exposition regarding a great legal and historical tradition of immigrants finding new homes in distant places, and how these new homes eventually came to supplant former homes.³⁶⁰ Attorneys

³⁵² *Quock Ting v. United States*, 140 U.S. 422, 425 (1891).

³⁵³ *Id.* at 422, 425.

³⁵⁴ *Fong Yue Ting v. United States*, 13 S.Ct. 1035 (1893).

³⁵⁵ *Id.* at 1016-17, 1030; *Chinese Deportation Act*, 27 Stat. 25 (1892).

³⁵⁶ *Fong Yue Ting*, 13 S.Ct. 1017 (referring to the *Chinese Deportation Act*, 27 Stat. 25, sec. 6 (1892)).

³⁵⁷ *Fong Yue Ting v. United States*, 13 S.Ct. 1030 (1893).

³⁵⁸ *Id.*

³⁵⁹ *Id.* at 1031.

³⁶⁰ *Id.* at 1032-33.

for Fong Yue Ting had argued that deportation constituted cruel and unusual punishment in violation of the Eighth Amendment,³⁶¹ and apparently Brewer wholeheartedly agreed. Brewer wrote, “But it needs no citation of authorities to support the proposition that deportation is punishment. Every one knows that to be forcibly taken away from home and family and friends and business and property, and sent across the ocean to a distant land, is punishment, and that oftentimes most severe and cruel.”³⁶² Brewer’s view would find resonance in later cases such as *Harisiades v. Shaughnessy* (1952), *Galvan v. Press* (1954), *Gastelum–Quinones v. Kennedy* (1963), and *U.S. v. Restrepo* (1992).³⁶³ More contemporaneously, his perspective found a kindred spirit in the remarkable Judge Learned Hand.

Law professor Geoffrey R. Stone described Learned Hand as the author of some 4,000 opinions and “the greatest judge of the twentieth century never to sit on the Supreme Court.”³⁶⁴ Hand’s lower court opinions were often cited by Supreme Court justices, and Hand himself was naturally very knowledgeable about the high court’s historical and recent decisions. He almost definitely knew of Brewer’s position regarding deportation, and sounded similar sentiments as early as *Klonis v. Davis* (1926) regarding a Polish immigrant. Walter Kronis came to the United States as a boy, ten years old at most, and unfortunately proceeded to live a life of crime that included two prison terms. It was a final crime involving “moral turpitude” that occasioned his deportation hearing. Fully acknowledging Kronis’s criminal past, Hand nonetheless wrote,

we think it not improper to say that deportation under the circumstances would be deplorable. Whether the relator came here in arms or at the age of ten, he is as much our product as though his mother had borne him on American soil. He knows no other language, no other people, no other habits, than ours; he will be as much a stranger in Poland as any one born of ancestors who immigrated in the seventeenth century. However heinous his crimes, deportation is to him exile, a dreadful punishment, abandoned by the common consent of all civilized peoples. Such, indeed, it would be to any one, but to one already proved to be incapable of honest living, a helpless waif in a strange land, it will be utter destruction. That our reasonable efforts to rid ourselves of unassimilable

³⁶¹ SALYER, *supra* note 56, at 49.

³⁶² *Fong Yue Ting*, 13 S.Ct. 1033.

³⁶³ See *Harisiades v. Shaughnessy*, 342 U.S. 600 (1952); *Galvan v. Press*, 347 U.S. 535 (1954); *Gastelum-Quinones v. Kennedy*, 374 U.S. 469, 480 (1963); *U.S. v. Restrepo*, 802 F.Supp. 791 (E.D.N.Y., 1992).

³⁶⁴ GEOFFREY R. STONE, *PERILOUS TIMES: FREE SPEECH IN WARTIME FROM THE SEDITION ACT OF 1798 TO THE WAR ON TERRORISM* 200 (2004). See also the editor’s comments in *LEARNED HAND, THE ART AND CRAFT OF JUDGING: THE DECISIONS OF JUDGE LEARNED HAND* 62 (Hershel Shanks, ed., Macmillan, 1968) as well as *JEROME FRANK, A MAN’S REACH: THE PHILOSOPHY OF JUDGE JEROME FRANK* 70 (1965).

immigrants should in execution be attended by such a cruel and barbarous result would be a national reproach.³⁶⁵

Hand maintained this viewpoint in later cases of the 1930s and 1940s.³⁶⁶ Brewer's precedent regarding deportation of Chinese is older than Hand's, of course, and took place amidst the nadir of nativism and in defiance of the 1892 Chinese Exclusion Act. Hand came later, after 1920s immigration restrictions assuaged the worst of the nativist frenzy. But obviously both were contributing to the same moral cosmopolitan tradition, and both continued to be cited throughout the twentieth century.

The prevailing racist society of the late nineteenth, early twentieth centuries did not render the Supreme Court unreasonable, even if the Court sent mixed signals. In 1892 the Supreme Court ordered the release of imprisoned Lau Ow Bew based upon habeas corpus rights.³⁶⁷ Justice Stephen Johnson Field, who spewed dogmatic racist ideology in *Chew Hoeng* (1884),³⁶⁸ nonetheless dissented along with Justice Brewer in *Fong Yue Ting* (1893)³⁶⁹ — even while remaining proud and unrepentant of his earlier decision. Field also dissented in part from the majority opinion in *Wong Wing v. U.S.* (1896), which involved a customs official sentencing four Chinese men to hard labor. The customs official concluded that the men were in the United States illegally, then utilized a provision in the 1892 Chinese Exclusion Act to sentence them to hard labor rather than immediate deportation.³⁷⁰ Field predictably upheld deportation of undocumented aliens, but denounced the government's failure to grant said aliens full constitutional and legal protection while temporarily in the United States.³⁷¹ These may seem like trivial objections on Field's part, given his adamant support of the 1882 Exclusion Act, but they illustrate the retention of important constitutional principles even on the part of an avowed racist.³⁷²

In 1898, despite dissents from Justice Fuller and Harlan, the majority of the Court remarkably upheld birthright citizenship in *U.S. v. Wong Kim Ark*, even when the parents were non-citizens.³⁷³ Wong Kim Ark was born in San Francisco in 1873. Customs officials detained him upon his return to California in 1895, after a visit to China, denying him entry and refus-

³⁶⁵ United States, *ex rel.* Klonis v. Davis, 13 F.2d 630, 631-32 (2d Cir. 1926). For later cases citing this passage, see United States, *ex rel.* Kowalenski v. Flynn, 17 F.2d 526 (D.C. Cir. 1927); Haller v. Esperdy, 397 F.2d 214 (2d Cir. 1968); Velez-Lozano v. Immigration and Naturalization Service, 463 F.2d 1309 (D.C. Cir. 1972).

³⁶⁶ See United States, *ex rel.* Guarino v. Uhl, 107 F.2d 399 (2d Cir. 1939); Di Pasquale v. Karnuth, 158 F.2d 880 (2d Cir. 1947).

³⁶⁷ Lau Ow Bew v. United States, 144 U.S. 59, 65 (1892).

³⁶⁸ Chew Heong v. United States, 112 U.S. 562-63, 569-70 (1884).

³⁶⁹ Fong Yue Ting v. U.S., 149 U.S. 747, 751, 761 (1893).

³⁷⁰ Wong Wing v. United States, 163 U.S. 230 (1896); Chinese Exclusion Act, 27 Stat. 25 § 4 (1882).

³⁷¹ *Wong Wing*, 163 U.S. 240, 244.

³⁷² For examples of Field's racism, see FRITZ, *supra* note 55, at 225, 237-38, 247, 248.

³⁷³ United States v. Wong Kim Ark, 169 U.S. 706 (1898).

ing to recognize his citizenship status.³⁷⁴ A lower court reversed this action and the United States government appealed, but the Supreme Court affirmed the lower court's ruling.³⁷⁵ This point of law stemmed from the 1866 Civil Rights Act,³⁷⁶ and was reaffirmed in the Fourteenth Amendment in 1868,³⁷⁷ but the Court had now removed any doubt that the law applied to people of Asian descent; an exceptional victory amidst the nativist era.

U.S. v. Sing Tuck (1904) involved five Chinese trying to enter the United States by way of Canada. There was some question about their American citizenship status, and Justice Oliver Wendell Holmes reversed an appellate court's order to investigate said status, stating that "A mere allegation of citizenship is not enough."³⁷⁸ In his dissent, Justice Brewer again sounded the moral cosmopolitan clarion cry:

The time has been when many young men from China came to our educational institutions to pursue their studies; when her commerce sought our shores, and her people came to build our railroads, and when China looked upon this country as her best friend. If all this be reversed and the most populous nation on earth becomes the great antagonist of this republic, the careful student of history will recall the words of Scripture, 'they have sown the wind, and they shall reap the whirlwind,' and for cause of such antagonism need look no further than the treatment accorded during the last twenty years by this country to the people of that nation.³⁷⁹

Justice Brewer had occasion to dissent again the following year in *U.S. v. Ju Toy*. Here the Court denied habeas corpus rights to Ju Toy, after the Secretary of Commerce and Labor denied him re-entry into the United States, despite Ju Toy claiming American citizenship.³⁸⁰ Justice Brewer found this "appalling" and cited numerous precedents to support his dissent.³⁸¹ He described the Court as banishing a citizen and stripping him of his rights in violation of congressional intent and constitutional protection.³⁸² Brewer's influence may have carried weight as soon as three years later in a somewhat similar case, *Chin Yow v. U.S.* (1908), when Justice Holmes' granted habeas corpus rights for citizenship determination. Where Holmes had stated in *Ju Toy* that due process did not require a judicial trial,³⁸³ in *Chin Yow* he seemed to reverse his previous stance by writing, "The courts must deal with the matter somehow, and there seems to be no

³⁷⁴ *Id.* at 650.

³⁷⁵ *Id.* at 706.

³⁷⁶ Civil Rights Act 1866, 14 Stat. 27-30, §1.

³⁷⁷ U.S. CONST., amend. XIV, §1.

³⁷⁸ *United States v. Sing Tuck*, 194 U.S.171 (1904).

³⁷⁹ *Id.* at 183.

³⁸⁰ *United States v. Ju Toy*, 198 U.S. 253 (1905).

³⁸¹ *Id.* at 270-74.

³⁸² *Id.* at 280-81.

³⁸³ *Id.* at 264.

way so convenient as a trial of the merits before the judge.”³⁸⁴ Brewer naturally concurred.

These seemingly small victories of the late nineteenth and early twentieth centuries nevertheless laid some important groundwork. In 1905 Holmes deferred to the Secretary of Commerce and Labor in *Ju Toy* regarding habeas corpus, but by 1920 Justice John Hessin Clarke was asserting judicial supremacy over the Secretary in the same matter.³⁸⁵ The nativist hysteria was dying down with the overall shift to circumscribed immigration. In this changing context, Clarke made the remarkable declaration that, “It is better that many Chinese immigrants should be improperly admitted than that one natural born citizen of the United States should be permanently excluded from his country.”³⁸⁶

For Asian American history, perhaps the most remarkable case of the 1920s was *Farrington v. Tokushige* (1927), wherein the Supreme Court upheld the right of Japanese Americans to educate their children in Japanese language schools.³⁸⁷ *Farrington v. Tokushige* constituted the third in a trio of remarkable victories against the nativist insistence upon WASP culture, the much-noted predecessors being *Meyer v. Nebraska* (1923)³⁸⁸ and *Pierce v. Society of Sisters* (1925).³⁸⁹ *Meyer* and *Pierce* were the pioneer cases against nativist reactions against white Catholic schools, but *Farrington* pushed the envelope further by extending Supreme Court jurisprudence into Asian language parochial schools.³⁹⁰ *Meyer* and *Pierce* dealt with ethnic groups of different languages and religion then found in mainstream WASP America, whereas *Farrington* more specifically dealt a blow (almost inadvertently) against prejudice against a different language and a different race. In *Farrington v. Tokushige*, Justice James Clark McReynolds wrote, “it would deprive parents of fair opportunity to procure for their children instruction which they think important and we cannot say is harmful. The Japanese parent has the right to direct the education of his own child without unreasonable restrictions; the Constitution protects him as well as those who speak another tongue.”³⁹¹ This was a remarkable multicultural inroad for the 1920s, but the Hawaiian context must be considered here, with its much larger proportional population of people of Asian descent. On the mainland, particularly in California and Washington, the discriminatory land laws alone reflected nativism’s enduring legacy.³⁹²

³⁸⁴ *Chin Yow v. United States*, 208 U.S. 14 (1908).

³⁸⁵ *Kwock Jan Fat v. White*, 253 U.S. 465 (1920).

³⁸⁶ *Id.*

³⁸⁷ *Farrington v. Tokushige*, 273 U.S. 284 (1927).

³⁸⁸ *Meyer v. Nebraska*, 262 U.S. 390 (1923).

³⁸⁹ *Pierce v. Soc’y of the Sisters of the Holy Names of Jesus*, 268 U.S. 510 (1925).

³⁹⁰ ROSS, *supra* note 340, at 5-6.

³⁹¹ *Farrington*, 273 U.S. at 299.

³⁹² See *Cockrill v. People*, 268 U.S. 258 (1925); *Frick v. Webb*, 263 U.S. 326 (1923); *Porterfield v. Webb*, 263 U.S. 225 (1923); *Terrace v. Thompson*, 263 U.S. 197 (1923); *Webb v. O’Brien*, 263 U.S. 313 (1923).

The Great Depression put a damper upon previous nativistic controversies, partly because immigration to the United States dramatically dropped in light of high unemployment rates and the absence of an economic incentive to immigrate.³⁹³ The next era of racist frenzy against Asians occurred after the Japanese bombing of Pearl Harbor, Hawaii, that initiated World War II for the United States. The Chinese were wartime allies with the United States against Japan and benefitted accordingly, notably beginning with the Chinese Exclusion Repeal Act the 1943.³⁹⁴ Yet even amidst this period, when anti-Japanese racists exploited the wartime emergency, the moral cosmopolitan dissent far from disappeared. In fact, in some ways it grew more vehement.

The notorious *Korematsu*³⁹⁵ and famous *Endo*³⁹⁶ decisions evoked strong responses from the legal community.³⁹⁷ Justice Frank Murphy's dissent in *Korematsu* may be his most famous, in which he called upholding the constitutionality of Japanese-American internment as falling into the "ugly abyss of racism."³⁹⁸ Murphy acknowledged the legitimacy of wartime emergencies as expressed in martial law, but ultimately saw the internment as an unconstitutional over-reaction and a dangerous civilian surrender to military authority.³⁹⁹ This last point alone is quite profound, but equally so was Murphy's adoption of the concept of racism, then a fairly new idea that arose in response to the Nazis' Aryan super race mythology.⁴⁰⁰ Murphy openly confronted the "questionable racial and sociological grounds"⁴⁰¹ that racists and nativists had actually used to rationalize their discrimination against Asians since the mid-nineteenth century. Murphy repeated his charges of racism in his concurrence with the unanimous decision in *Endo*,⁴⁰² which ruled to release one Mitsuye Endo from wartime detention.⁴⁰³

³⁹³ William S. Bernard, *A History of U.S. Immigration Policy*, in IMMIGRATION, at 99 (1982).

³⁹⁴ 78 Cong. Ch. 344, Dec. 17, 1943, 57 Stat. 600.

³⁹⁵ *Korematsu v. United States*, 323 U.S. 214 (1944).

³⁹⁶ *Ex parte Endo*, 323 U.S. 283 (1944).

³⁹⁷ For examples, see Maurice Alexander, *Wartime Control of Japanese-Americans*, 28 CORNELL L.Q. 385, 1942-1943; Robert E. Cushman, *Some Constitutional Problems of Civil Liberty*, 23 B.U. L. REV. 335,(1943); Nanette Dembitz, *Racial Discrimination and the Military Judgment: The Supreme Court's Korematsu and Endo Decisions*, 45 COLUM. L. REV. 175 1945; MILTON R. KONVITZ, THE ALIEN AND THE ASIATIC IN AMERICAN LAW 277 (1946); Eugene V. Rostow, *The Japanese American Cases—A Disaster*, 54 YALE L.J. 489, (1944-1945).

³⁹⁸ *Korematsu*, 323 U.S. at 234.

³⁹⁹ *Id.* at 235-37.

⁴⁰⁰ GEORGE M. FREDRICKSON, RACISM: A SHORT HISTORY 5 (2002).

⁴⁰¹ *Korematsu*, 323 U.S. at 237.

⁴⁰² *Endo*, 323 U.S. at 308.

⁴⁰³ *Id.* at 305.

In 1942, Mitsuye Endo petitioned for habeas corpus after being forcibly relocated to the Tule Lake internment camp.⁴⁰⁴ The basis for her habeas corpus claim was her loyalty as a United States citizen, absence of any criminal record, and her detainment against her will. The Department of Justice and even the War Relocation Authority concurred in all these claims and the Supreme Court agreed, granting habeas corpus and ruling that Endo was entitled to release. Justice William O. Douglas, writing the Court's majority opinion, noted, "A citizen who is concededly loyal presents no problem of espionage or sabotage. Loyalty is a matter of the heart and mind not of race, creed, or color."⁴⁰⁵ Even amidst the war, the moral cosmopolitan cause was gaining some ground.

Colorado Governor Ralph Carr (1939-43) significantly protested the internment of Japanese-Americans in his state's federally-designated Granada Relocation Camp, claiming that such internment was a violation of civil rights.⁴⁰⁶ Colorado voters must have sympathized. Anti-Japanese wartime hysteria hit Colorado in 1944 with the state house proposing a constitutional amendment prohibiting landownership by people of Japanese descent. Colorado voters, however, decided against the amendment by a significant margin, 184,458 to 168,865.⁴⁰⁷ Douglas R. Hurt interpreted this vote as a reflection of both high-minded anti-racism as well as pragmatism, for denying one group of immigrants or their descendants land ownership rights would have ominous future implications for other immigrants and their descendants.⁴⁰⁸ Historically, ethnic minorities and their mainstream allies have had to organize and finance their arguments for establishing these new social norms,⁴⁰⁹ and this Colorado instance was a stellar example of such an effort.⁴¹⁰ Although nativist and racist elements will always remain, after the 1920s nativists either began to take assurance with new laws limiting immigration or, as William G. Ross argues, slowly recognized that their WASPish values and ethnic minorities need not necessarily exclude the other,⁴¹¹ or perhaps some combination of both and other factors.

Even during the World War II era, and definitely afterward, the United States began to take a leadership role in international race relations,

⁴⁰⁴ *Id.* at 285-86, 295.

⁴⁰⁵ *Id.* at 303.

⁴⁰⁶ *Ralph Carr Award to Honor Japanese-Americans Supporters*, DAILY YOMIURI ONLINE (July 18, 2012), available at <http://www.yomiuri.co.jp/dy/world/T120717003926.htm>; THE GOVERNOR RALPH L. CARR COLLECTION at the *Colorado State Archives*, available at <http://www.colorado.gov/dpa/doit/archives/govs/carr.html>.

⁴⁰⁷ BILL HOSOKAWA, COLORADO'S JAPANESE AMERICANS: FROM 1886 TO THE PRESENT 118 (2005); DOUGLAS R. HURT, THE GREAT PLAINS DURING WORLD WAR II 300-301 (2008).

⁴⁰⁸ HURT, *supra* note 407, at 300, 301.

⁴⁰⁹ ROSS, *supra* note 340, at 205-06; Webber, *supra* note 340, at 88.

⁴¹⁰ HOSOKAWA, *supra* note 407, at 118.

⁴¹¹ ROSS, *supra* note 340, at 205.

particularly regarding upholding the United Nations charter.⁴¹² The booming economy of the 1950s also helped mitigate employment competition antagonism associated with immigrants. Certainly the moral cosmopolitan precedent was not the only factor in the amelioration of anti-Asianism and rectification for past nativist and racist wrongs. So this is not to advocate a determinist argument, in which the moral cosmopolitans were “leading up to” what later happened. Instead, the moral cosmopolitans set examples of an alternative social reality that placed it well within the scope of Enlightenment principles, and thus hardly representative of a radical fringe, then or now. For a variety of reasons, including the United States’ post-war role in the United Nations’ advocacy of universalist human rights, American society gradually shifted away from institutionalized nativism and racism to embrace what their moral cosmopolitan predecessors had exemplified during the nineteenth and early twentieth centuries. It would be impossible to make a scientific cause-and-effect connection between the moral cosmopolitan precedent and the later liberalization of immigration law, and yet the change in social, legislative, and judicial mood must have some basis in the past. Large historical movements never come out of a vacuum. Christianity had centuries of cultural preparation in Judaism and the Eastern Mystery religions.⁴¹³ Marxism had many precursors among French “utopian” socialists.⁴¹⁴ The American Civil Rights era of the 1950s-1960s, in fact, had decades of preparation and pump-priming in the efforts of the NAACP’s Legal Defense Fund and in the work of Charles Hamilton Houston at Howard Law School alone.⁴¹⁵ The current study is no exception. The post-war liberalization of immigration law and improved social, economic, educational, and political advancement for Asians and Asian-Americans had a small but crucial precedent prior to and in the middle of the most dire forces against them.

The moral cosmopolitan precedent came to fruition during the post-war era. Liberalized immigration laws particularly benefited Asian peoples, as did the general sweep of the Civil Rights Movement and Affirmative

⁴¹² FORD, *Rights Gone Wrong*, *supra* note 340, at 127-29, 158 (2012); *Oyama v. California*, 332 U.S. 633, 674 (1948); *Sei Fujii v. State*, 242 P.2d 617, 721, 725, 742 (Cal. 1952).

⁴¹³ *See generally* SAMUEL ANGUS, *THE MYSTERY-RELIGIONS AND CHRISTIANITY: A STUDY IN THE RELIGIOUS BACKGROUND OF EARLY CHRISTIANITY* (1925); RICHARD REITZENSTEIN, *HELLENISTIC MYSTERY-RELIGIONS: THEIR BASIC IDEAS AND SIGNIFICANCE* (John E. Steely trans., 1978).

⁴¹⁴ A study of a specific utopian socialist may be found in JONATHAN BEECHER, *CHARLES FOURIER: THE VISIONARY AND HIS WORLD* (1986). *See generally* KEITH TAYLOR, *THE POLITICAL IDEAS OF THE UTOPIAN SOCIALISTS* (1982).

⁴¹⁵ RICHARD KLUGER, *SIMPLE JUSTICE: THE HISTORY OF BROWN V. BOARD OF EDUCATION AND BLACK AMERICA’S STRUGGLE FOR EQUALITY* 126 (1976); GENNA RAE MCNEIL, *GROUNDWORK: CHARLES HAMILTON HOUSTON AND THE STRUGGLE FOR CIVIL RIGHTS* Ch. 6 (1983); MARK V. TUSHNET, *MAKING CIVIL RIGHTS LAW: THURGOOD MARSHALL AND THE SUPREME COURT, 1936-1961* 6,7 (1994); Leland Ware, *Charles Houston’s Transformation of Legal Education*, 32 *HOWARD L. REV.* 479, 479-93 (1989).

Action. The aforementioned 1943 Chinese Exclusion Repeal Act⁴¹⁶ began the liberalization of immigration, albeit with extreme modesty. In *Oyama v. California* (1948)⁴¹⁷ the Supreme Court finally rejected alien land law, and the West Coast states eventually followed this precedent.⁴¹⁸ A very modest liberalization of the immigration quota for Japanese arrived with the 1952 McCarran-Walter Act.⁴¹⁹ The crowning legislation of liberalized immigration arrived at the apex of the Civil Rights Movement with the Immigration and Nationality Act of 1965.⁴²⁰ The Civil Liberties Act of 1988 tried to make amends and offered reparations for the wartime Japanese internment.⁴²¹ There were many other pieces of legislation that fit into this general civil rights revolutionary context,⁴²² but by the mid-1960s law and society were clearly beginning to reflect a sea change in America that the earlier generations of moral cosmopolitans had favored all along.

All together, here we might examine how moral cosmopolitanism of the pre- and post-Civil Rights eras came to mesh with old and new ideas regarding “melting pot” America. Describing the contemporary socio-political climate, James E. Bond delineates a useful dichotomy between “Melting Pot” multiculturalists and “Salad Bowl” multiculturalists.⁴²³ Melting Pot multiculturalists see cultural assimilation as central to the American vision, with numerous reasonable provisos for accommodating cultural difference.⁴²⁴ Salad Bowl multiculturalists, on the other hand, are the post-Civil Rights Era celebrators of cultural difference, who have ironi-

⁴¹⁶ Chinese Exclusion Repeal Act, 78 Cong. Ch. 344, Dec. 17, 1943, 57 Stat. 600 (1943).

⁴¹⁷ *Oyama*, 332 U.S. 633.

⁴¹⁸ *Kenji Namba v. McCourt*, 204 P.2d 569 (Or. 1949); *Sei Fujii*, 242 P.2d 617; Thomas E. Stuen, *Asian Americans and Their Rights for Land Ownership*, in HYUNG-CHAN KIM, ASIAN AMERICANS AND THE SUPREME COURT 620 (1992).

⁴¹⁹ Formally known as the Immigration and Nationality Act, Pub.L. 82-414, 66 Stat. 163, 82nd Congress (June 27, 1952).

⁴²⁰ Immigration and Nationality Act, Pub.L. 89-236; 79 Stat. 911, 89th Congress (Oct. 3, 1965).

⁴²¹ The Civil Liberties Act, Pub.L. 100-383, 102 Stat. 903 (Aug. 10, 1988).

⁴²² See e.g., *Takahashi v. Fish & Game Comm'n* 334 U.S. 410 (1948). Other examples affecting Asians included the War Brides Act, Pub. L. 79-271; 59 stat 659, 79th Cong. (Dec. 28, 1945) and Alien Fiancées and Fiancés Act, Pub. L. 79-471; 60 stat. 339, 79th Cong. (June 29, 1946). Huping Ling is one of the few scholars to acknowledge the importance of these comparatively minor changes in law. See HUPING LING, CHINESE CHICAGO: RACE, TRANSNATIONAL MIGRATION, AND COMMUNITY SINCE 1870 205-07 (2012). A list that places all these pieces of legislation within a chronological context may be found at

<http://library.uwb.edu/guides/usimmigration/USimmigrationlegislation.html>

⁴²³ James E. Bond, *Multiculturalism: America's Enduring Challenge*, SEATTLE J. FOR SOC. JUSTICE 1, 59-70 (Spring/Summer, 2002).

⁴²⁴ *Id.* at 60-61.

cally re-segregated American society along ethnic lines,⁴²⁵ with varying degrees of cultural authenticity.⁴²⁶

Bond's dichotomy has profound implications that partly coalesce with notions of nationalism ("patriots of soil") and patriotism ("patriots of liberty").⁴²⁷ I'm using the term "nationalists" to refer to the exclusionary movements that arose during the nineteenth and twentieth centuries, particularly associated with Germany, but clearly manifested in the United States and elsewhere.⁴²⁸ This was, perhaps, the darkest aspect of the Romantic period, which otherwise offered many valuable artistic and cultural contributions and aspirations. So "patriots," in this context, would then mean loyalty to Enlightenment era ideals of universal human rights, moral cosmopolitanism, tolerance, and acceptance of human differences⁴²⁹—hence the concept of "patriots of liberty."

In one of our greatest post-war ironies, Salad Bowl multiculturalists and nationalists have ended up with a common bigotry that dismisses Melting Pot multiculturalism, the latter embodying a patriotism of liberty, with favorable moral cosmopolitan implications that have always included interracial benevolence. The latter is the historic focus of this study, which contemporary Salad Bowl multiculturalists have concertedly dismissed.

⁴²⁵ *Id.* at 61.

⁴²⁶ The latter point involves self-conscious culture, usually for opportunistic ends. See BRIAN BARRY, CULTURE AND EQUALITY: AN EGALITARIAN CRITIQUE OF MULTICULTURALISM 259 (2002); FORD, RACIAL CULTURE, *supra* note 5, at 35, 39, 41, 44, 61, 134; RICHARD THOMPSON FORD, RACE CARD: HOW BLUFFING ABOUT BIAS MAKES RACE RELATIONS WORSE 192, 278 (2008). This phenomenon certainly occurs outside the United States. See generally THE INVENTION OF TRADITION (Eric Hobsbawm & Terence Ranger eds., 1992).

⁴²⁷ Here I follow MAURIZIO VIROLI, FOR LOVE OF COUNTRY: AN ESSAY ON PATRIOTISM AND NATIONALISM (1995). "Patriots of soil" is a term from Enlightenment era philosopher, Earl of Shaftesbury. See ANTHONY ASHLEY COOPER, EARL OF SHAFTESBURY, CHARACTERISTICS OF MEN, MANNERS, OPINIONS, TIMES 403 (Lawrence E. Klein, ed., Cambridge Univ. Press, 1999) (1711). "Patriots of liberty" is Viroli's term, implicated in contrast. See VIROLI at 59, 184-85.

⁴²⁸ BARRY, *supra* note 426, at 308-09; Dorothy E. Roberts, *Why Culture Matters to Law: The Difference Politics Makes*, in CULTURAL PLURALISM, IDENTITY POLITICS, AND THE LAW 102 (Austin Sarat & Thomas R. Kearns eds., 1999). For a contemporary reaction against German nationalism in particular, but not uncritical of Britain and the United States, see CHARLES WALDSTEIN, PATRIOTISM, NATIONAL AND INTERNATIONAL: AN ESSAY (1917).

⁴²⁹ VIROLI, *supra* note 427, at 2, 8-9, 57, 58, 184-85; WALDSTEIN, *supra* note 428, at 145, 146, 149.

their own assimilation.⁴³⁴ In the broadest sense, however, minority cultures have constantly contributed to the ever-evolutionary mainstream. This “enrichment assimilation” is a salient feature in a land of immigrants such as the United States, consistent with liberty patriotism. Mainstream American culture today barely resembles mainstream culture of previous decades. For a moral cosmopolitan advocate, this is what makes mainstream American culture so wonderful. Some of the fringe cultures are interesting; some frightening. But the self-consciously manufactured “minority cultures” are mainly the tools of opportunists who have to invent something that supposedly distinguishes them.

Outside of academia, many people would find enrichment assimilation merely a matter of common sense. The late newspaper columnist, writer and historian Bill Hosokawa offered a sober antidote for the Salad Bowl perspective. Hosokawa unabashedly attributed the great success of Japanese-Americans to a mixture of both mainstream American cultural traits as well as those rooted in the old country, such as filial piety, a strong sense of family and family honor (and, by extension, a strong sense of community) as well an attitude of humility, duty, and making the most of difficult circumstances.⁴³⁵ “Isn’t the melding of cultures what America is all about?” Hosokawa asked.⁴³⁶ No complaining about the model minority stereotype here. Instead, Hosokawa praised the “happy combination of the more admirable of Japanese traits being nurtured in the freedom and openness of American society” as responsible for stellar success of Japanese-Americans.⁴³⁷ Sam Chan expressed a similar pro-assimilation sentiment in a 1943 letter to his congressman, Representative B. Carroll Reece of Tennessee. Chan wrote, “I dislike to boast, but our people in America are loyal, honest, and obedient to your laws. No other race can show such an enviable record.”⁴³⁸ As Huping Ling has demonstrated, there is a deep tradition among Chinese immigrants to America for promoting assimilation into the mainstream through “Americanization” or at least “hybrid” social organizations that helped bridge the two cultures.⁴³⁹ Chinese intellectuals in particular saw themselves as cultural liaisons or cultural agents whose duty it was to help bridge the two cultures.⁴⁴⁰

⁴³⁴ FRIEDMAN, *AMERICAN LAW*, *supra* note 15, at 343, 545; DAVID M. KENNEDY, *FREEDOM FROM FEAR: THE AMERICAN PEOPLE IN DEPRESSION AND WAR, 1929-1945* 306 (1999); JAMES S. OLSON, *EQUALITY DEFERRED: RACE, ETHNICITY, AND IMMIGRATION IN AMERICA, SINCE 1945*, 11, 144-49 (2002).

⁴³⁵ HOSOKAWA, *supra* note 407, at 247, 248, 250.

⁴³⁶ *Id.* at 250.

⁴³⁷ *Id.*

⁴³⁸ Sam Chan to Rep. B. Carroll Reece, June 8, 1943, 78th Cong., 1st sess., *Appendix to the Congressional Record* at A2857.

⁴³⁹ LING, *supra* note 422, at 38-42, 149, 163.

⁴⁴⁰ *Id.* at 172, 173, 189.

VI. LIBERTY PATRIOTISM IMPLICATIONS

Soil patriots rally around the flag and ignore the ambiguity and potential shallowness of its symbolism. Interviews with various people about the meaning of the Confederate flag illustrate this point quite vividly. Liberty patriots, on the other hand, rally around constitutionalism. They fully recognize the historic mistakes of constitutionalism, both in its formal and populist manifestations. They see constitutionalism as an eternal arena for debate and disagreement, but also as one of the nation's greatest tools for aspiring toward future ideals. To state the obvious, this is particularly important for moral cosmopolitanism to survive in a nation of immigrants and such cultural variety.

Tribal or clan units are among our oldest social structures.⁴⁴¹ Early people, like other mammal groups (lions, wolves, horses, and all primates) banded together out of necessity for survival.⁴⁴² Native American tribes often had names for themselves that simply meant "the people,"⁴⁴³ meaning us, our group. Group security or insecurity centered around the gathering of and competition for resources. The tribal punishment of ostracism,⁴⁴⁴ sometimes equivalent with a death sentence since solitary survival was almost impossible, illustrated the profound group-orientation of early people.⁴⁴⁵ After the historical era arrived, people continued affiliations to clans and developed all sorts of orientations to geographical neighbor-

⁴⁴¹ KENNETH S. CARLSTON, *SOCIAL THEORY AND AFRICAN TRIBAL ORGANIZATION: THE DEVELOPMENT OF SOCIO-LEGAL THEORY* 3, 15, 16, 65 (1968).

⁴⁴² See e.g. JAN VANSINA, *HOW SOCIETIES ARE BORN: GOVERNANCE IN WEST CENTRAL AFRICA BEFORE 1600* 26-33 (2005).

⁴⁴³ Phil Konstantin has compiled a long list of such tribal names. See Phil Konstantin, *Tribal Name Meanings and Alternative Names*, available at <http://americanindian.net/names.html>.

⁴⁴⁴ An evocative depiction of tribal ostracism may be found toward the end of the remarkable, multiple award winning film, *FAST RUNNER* (2002), shot among the Canadian Inuit. For another aboriginal example pertaining to the Cheyenne, see KARL N. LEWELLYN & E. ADAMSON HOEBEL, *The Tribal Ostracism and Reinstatement of Sticks Everything Under His Belt*, in *THE CHEYENNE WAY* 145, 154-7 (1941). For an Afghanistan tribal example, see Jiloufer Qasim Mahdi, *Pukhtunwali: Ostracism and Honor Among the Pathan Hill Tribes*, 7 *ETHOLOGY AND SOCIOBIOLOGY* 295 (1986); see also Patrice H. Kunesh, *Banishment as Cultural Justice in Contemporary Tribal Legal Systems*, 37 *N.M. L. REV.* 85 (2007); Eric A. Posner, *The Regulation of Groups: The Influence of Legal and Nonlegal Sanctions on Collective Action*, 63 *U. CHI. L. REV.* 133 (1996). Unfortunately, the much-publicized use of "traditional" Tlingit ostracism to punish two Tlingit boys by banishing them to a remote southeastern Alaska island apparently turned out to be fraudulent on at least two levels: Rudy James falsely represented himself as a Tlingit tribal judge and Tlingit culture has no tradition of ostracism. See Timothy Egan, *Indian Boys' Exile Turns Out to be a Hoax*, *N.Y. TIMES* Aug. 31, 1994.

⁴⁴⁵ After the Greeks abandoned strictly tribal orientation and began developing democratic politics, exile and ostracism became a more institutional practice (later apparently influencing the Romans). See generally SARA FORSDYKE, *EXILE, OSTRACISM, AND DEMOCRACY: THE POLITICS OF EXPULSION IN ANCIENT GREECE* (2005).

hoods,⁴⁴⁶ religions, economic systems, military rivalries, kingdoms, and empires. Nationalism or soil patriotism is a comparatively new phenomenon, not beginning to develop fully until the nineteenth century.⁴⁴⁷ Particularly with the post-World War II formation of organizations such as the United Nations,⁴⁴⁸ UNESCO, the World Bank, Amnesty International, Environmental Law Alliance Worldwide, Doctors without Borders (and many other organizations) we have more fully entered a trans-national or supra-national era. Moral cosmopolitans may embrace trans-nationalism, but generally do not advocate ending localism or nationalism.⁴⁴⁹ The point is, “tribalism” in a general sense is part of the human condition, yet obviously we are fully capable of aspiring toward and even realizing the greater ideal of understanding the universality of humanity. The United States would seem to have been a natural laboratory for fostering moral cosmopolitanism in a polyglot nation. But, advocacy and idealism aside, provincialism continues and will continue to manifest itself in soil patriotism mold.⁴⁵⁰ Still, the liberty patriotism alternative has been demonstrated, and continues to be demonstrated, by the moral cosmopolitan disposition.

Imperialism abroad and domestic racism and nativism live on in the United States, of course. If Indian tribal names mean “the People,” a great many Americans continue to consider themselves the chosen people, still showing the world what it is all about from their city upon the hill.⁴⁵¹ But moral cosmopolitans can nevertheless serve their local tribe through education, example, and by discouraging racism, bigotry, and hyper-nationalism.

⁴⁴⁶ CARLSTON, *supra* note 442, at 15, 16. Carlston deals with the rise of the state at 27-28.

⁴⁴⁷ A short list of the voluminous literature about nationalism might include FRANK COPPA, *THE ORIGINS OF THE ITALIAN WARS OF INDEPENDENCE* (1992); DENNIS E. SHOWALTER, *THE WARS OF GERMAN UNIFICATION* (2004); STANFORD J. SHAW, *HISTORY OF THE OTTOMAN EMPIRE AND MODERN TURKEY, v2: REFORM, REVOLUTION, AND REPUBLIC: THE RISE OF MODERN TURKEY, 1808-1975* (1977); JAMES M. MCPHERSON, *BATTLE CRY OF FREEDOM: THE CIVIL WAR ERA* (1988).

⁴⁴⁸ For the beginnings of the United Nations, see STEPHEN C. SCHLESINGER, *ACT OF CREATION: THE FOUNDING OF THE UNITED NATIONS, A STORY OF SUPERPOWERS, SECRET AGENTS, WARTIME ALLIES AND ENEMIES, AND THEIR QUEST FOR A PEACEFUL WORLD* (2003). For a more recent assessment (false dichotomy of the title notwithstanding), see *IRRELEVANT OR INDISPENSABLE?: THE UNITED NATIONS IN THE TWENTY-FIRST CENTURY* (Paul Heinbecker & Patricia Goff eds., 2005).

⁴⁴⁹ Kleingeld & Brown, *supra* note 8, at §§ 3.0, 3.3.

⁴⁵⁰ FORD, *RACIAL CULTURE*, *supra* note 5, at 164-65.

⁴⁵¹ John Winthrop’s famous 1630 phrase, “city upon a hill,” is often cited as the earliest example of Americans’ Calvinistic self-conception of exceptionalism. This cultural trait reemerged with industrial force during the 19th century under the banner of Manifest Destiny, in which the deity had ordained that Americans overtake what became the continental United States, and which then spilled into the Pacific with the domination of Hawaii, Guam, the Philippines, and other islands. See FREDERICK MERK, *MANIFEST DESTINY AND MISSION IN AMERICAN HISTORY: A REINTERPRETATION* (1963). For John Winthrop’s sermon, see *A Modell of Christian Charity* (1630), available at <http://history.hanover.edu/texts/winthmod.html>.

This essay has attempted to make the case for the historic significance of one, loosely-constructed such group.

In the most tolerant and most broadminded sense, perhaps we can forgive the people of the nineteenth century for their notions of race, racial hierarchy, and race supremacy that all seem ludicrously disturbing to us now. They were just discovering the wider world and that was their cosmopolitan moment. Already they had outgrown the Greeks' ancient concept of cosmopolitanism,⁴⁵² if for no other reasons than the advancement of transportation, communication, and cartography. Some remained mired in a contradictory Kantian cosmopolitanism which claimed to embrace the world while nevertheless allocating highest status for one's own tribe.⁴⁵³ But now it is our cosmopolitan moment. Many feel that cosmopolitanism is now "an urgent moral necessity."⁴⁵⁴ Will we embrace the best virtues of cultural relativism that cosmopolitanism necessitates? Or will we continue believing the self-righteous parochial myth of the United States as a City Upon a Hill?

VII. CONCLUSION

If racism derives from ignorance, fear, and hatred, then non-racism must derive from an absence of these negative qualities. Moreover, racists perceive and definitely attempt to impose inequality upon the targets of their discrimination.⁴⁵⁵ Friendship, as Aristotle and other philosophers have recognized, requires equality.⁴⁵⁶ Recognizing the place of interracial friendship among Asians and Caucasians amidst a harsh era of prevailing malevolence is an important, if largely ignored, aspect of (Asian) American history. It adds a vital dimension to a subject perhaps rightfully preoccupied, until recent years, with injustice.

Legal dissent is most obviously significant when a later judicial decision relies upon it to overturn the *status quo* and thus advance a new line of jurisprudence. Within the theoretical framework of populist constitutionalism, the moral cosmopolitans established a populist dissent that re-

⁴⁵² Kleingeld & Brown, *supra* note 8, at § 1.1.

⁴⁵³ Fine & Cohen, *supra* note 5, at 145.

⁴⁵⁴ *Id.* at 162. Also see Kleingeld & Brown, *supra* note 8, at § 3.3; Waldron, *supra* note 5, at 754, 777, 778.

⁴⁵⁵ Although employing the fallacy of the lonely fact (FISCHER, *supra* note 131, at 109-10), Kwame Appiah makes an important point regarding the dubious links between ignorance and racism by citing the example of the famous world traveler, Sir Richard Burton (1821-1890). Despite Burton's unusually high exposure and experience with numerous ethnic groups around the world, he nevertheless remained prejudiced against the non-English in the almost stereotypical Anglophile fashion. See APPIAH, *supra* note 5, at 1-8. However, Appiah's point does reinforce the moral cosmopolitan idea that an *attitude* or disposition toward embracing other peoples and cultures on their own terms (mainly regardless of actual physical proximity) is what constitutes a solution to provincialism, parochialism, and the prejudice that all too often accompanies such worldviews.

⁴⁵⁶ NICOMACHEAN ETHICS, *supra* note 139, at 214, 216, 217 Marilyn Friedman, *Friendship and Moral Growth*, 23 J. OF VALUE INQUIRY 1, 3 (1989).

mained largely sublimated in the national consciousness until the 1940s and afterward. This is not to argue for determinism in historical causation, nor to commit the presentist fallacy by reading modern values into the past.⁴⁵⁷ But the fact remains that the moral cosmopolitans were part of enormous (if sometimes amorphous) evolutionary forces that witnessed the gradual interpretation of “all men” created equal meaning landowning white males to then, quite literally, all peoples or at least all citizens and citizen aspirants. The confines of patriotism must remain, and perhaps they always will; but at least it can be a patriotism of liberty. The seeds of universalism were there in the Enlightenment, but obviously it has taken a great deal of time for them to reach their contemporary degree of partial fruition. To say that the moral cosmopolitans were ahead of their time commits the aforementioned historical causation problems. To say they and we share the same values commits the universalist fallacy.⁴⁵⁸ Instead, the moral cosmopolitans had an ultimately incalculable array of both personal and practical inclinations for defying the predominant racism and nativism of their era. They were braver and lonelier than contemporary cosmopolitan Americans. They were unselfconscious in their benevolence, which they expressed naturally without the influences of political correctness and the absence of even a shadow of affirmative action. They had a conception of humanity that was unusually generous for these earlier time periods. They knew the United States was a land overwhelmingly populated by immigrants, but disagreed that expansionist immigration would harm the country. It is difficult not to view them as heroic people.

⁴⁵⁷ FISCHER, *supra* note 131, at 135-40, 166-67, 183-86.

⁴⁵⁸ *Id.* at 203-07.

THE EXPANDING RIGHT TO AN EFFECTIVE REMEDY: COMMON DEVELOPMENTS AT THE HUMAN RIGHTS COMMITTEE AND THE INTER-AMERICAN COURT

Valeska David*

Ghent University

ABSTRACT

The right to an effective remedy –crystallised in numerous treaties– has evolved to comprise individual rights and States obligations of a complex nature. This article discusses the procedural and substantive implications of an expansive interpretation of this right by the Human Rights Committee (HRC) and Inter-American Court of Human Rights (IACtHR). The HRC’s case law has significantly influenced the way the IACtHR has conceived of a set of rights belonging to victims of gross human rights violations, including a right to access justice and to demand investigation, prosecution, punishment and truth. Notwithstanding the greater protection and participation this construction of the right offers to victims, its difficulties warrant a critical appraisal. Some of these difficulties are related to how the IACtHR endorsed the HRC’s jurisprudence despite the differences between the norms and practices of both organs. From another perspective, the right to an effective remedy within the HRC jurisprudence also covers the award of reparations. In this area the article discusses the possible influence of the IACtHR’s jurisprudence upon the HRC.

CONTENTS

I.	INTRODUCTION	260
II.	REMEDIES AND THE RIGHT TO AN EFFECTIVE REMEDY UNDER THE ICCPR	263
III.	THE EXPANSION OF A RIGHT TO A PROCEDURAL REMEDY IN THE HRC’S CASE LAW	264
	A. THE STATE DUTY TO INVESTIGATE, PROSECUTE AND PUNISH	265
	B. THE RIGHT TO TRUTH	268
IV.	THE INFLUENCE OF THE HRC’S EXPANSIVE INTER- PRETATION OF THE RIGHT TO AN EFFECTIVE REMEDY IN THE JURISPRUDENCE OF THE INTER- AMERICAN COURT	270

A. THE RIGHT TO AN EFFECTIVE REMEDY UNDER THE ACHR	270
B. THE STATE DUTY TO INVESTIGATE, PROSECUTE AND PUNISH.....	273
C. THE RIGHT TO TRUTH.....	277
V. THE EXPANSION OF THE SUBSTANTIVE NOTION OF A REMEDY: REPARATIONS IN THE HRC AND THE POSSIBLE IMPACT OF THE IACTHR'S JURISPRUDENCE.	279
A. REPARATIONS AT THE HRC.....	280
B. THE IACTHR'S REPARATIONS AS A SOURCE OF INSPIRATION	283
VI. CONCLUSIONS	284

I. INTRODUCTION

The right to an effective remedy is considered to be one of the most fundamental guarantees for the protection of human rights,¹ and a cornerstone in achieving justice for victims.² Pursuant to this right, States have an obligation to provide an effective remedy to all persons in their jurisdiction who assert an arguable claim that their rights, which the State has the primary responsibility to guarantee, have been violated.³ In other words, an effective remedy implies the possibility to make rights

*Attorney at Law, PhD researcher, Ghent University. LL.M. (*Cum Laude*), Utrecht University. Email: Valeska.david@ugent.be.

I would like to thank Brianne McGonigle for her insights and support of my work on this paper. All errors are my own. This research was completed thanks to the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office, more specifically the IAP "The Global Challenge of Human Rights Integration: Towards a Users' Perspective" (www.hrintegration.be).

¹ INTERNATIONAL COMMISSION OF JURISTS, *THE RIGHT TO A REMEDY AND TO REPARATIONS FOR GROSS HUMAN RIGHTS VIOLATIONS. PRACTITIONER'S GUIDE* No. 2, 44 (2006); *Castillo-Páez v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 34, ¶ 82, (Nov. 3, 1997); *Blake v. Guatemala*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 36, ¶ 102, (Jan. 24, 1998).

² CHERIF M. BASSIOUNI, *International Recognition of Victims Rights*, in 6 *HUM. RTS. L. REV.*, 203, 231 (2006).

³ Optional Protocol to the International Covenant of Civil and Political Rights, Dec. 16, 1966, S. Exec. Rep. 102-23, 999 U.N.T.S. 171, art. 5.2.b [hereinafter, Protocol I to the Geneva Conventions]; American Convention on Human Rights, Nov. 22, 1969, O.A.S. T.S. No. 36, 1144 U.N.T.S. 143, art. 46.2.a [hereinafter ACHR]; Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, 213 U.N.T.S. 221, art. 35.1 [hereinafter ECHR].

enforceable against a duty-bearer and, therefore, meaningful in practice. This explains the assertion that there is no right without a remedy.⁴

The right to an effective remedy is enshrined in global and regional human rights instruments,⁵ and its importance has led to it being recognized as non-derogable in character, particularly in relation to remedies for violations of those rights that cannot be suspended in a state of emergency.⁶ Moreover, through the interpretation and application of this right by certain international bodies for the protection of human rights, the content of the right to an effective remedy has expanded significantly. Under the International Covenant on Civil and Political Rights (hereinafter ICCPR) the right to an effective remedy has been broadly understood both from a procedural and a substantive point of view. The expansion of the scope of this right is related to the international recognition of a so-called duty to investigate, to prosecute, to punish, and the right to know the truth of gross human rights violations. The Human Rights Committee (hereinafter HRC), as the interpreter of rights with universal applicability, has played a key role in shaping this right and its multiple corresponding obligations beyond the United Nations system. At the regional level, the experience of the Inter-American Court of Human Rights (hereinafter IACtHR) constitutes an interesting example of the complexities that result from importing the expansive HRC construction of the right to an effective remedy despite the distinctive features that explain such an expansive construction.

This article argues firstly, that the right to an effective remedy has been developed extensively by the HRC in its case law, particularly in relation to grave human rights violations, and that this process has not been entirely consistent or free of difficulties. Secondly, the interpretation of the right by the HRC has significantly influenced the jurisprudence of the IACtHR, which changed its previous understanding of the right to an effective remedy, with some problematic results. Thirdly, the HRC's expansion of the right to an effective remedy is not limited to its procedural dimension; it has also extended its substantive dimension. The right has been used by the HRC as a legal basis upon which to request States to make reparation for the rights violated, in the absence of an express

⁴ MANFRED NOWAK, *Eight Reasons Why We Need a World Court of Human Rights*, in INTERNATIONAL HUMAN RIGHTS MONITORING MECHANISMS: ESSAYS IN HONOUR OF JAKOB TH. MÖLLER 697, (Gudmundur Alfredsson et al. ed. 2009); DINAH SHELTON, REMEDIES IN INTERNATIONAL HUMAN RIGHTS LAW 8 (2 ed. 2005).

⁵ Universal Declaration of Human Rights, G.A. res. 217A (III), U.N. Doc A/810 at 71, Dec. 10, 1948, art. 8; ACHR *supra* note 3, art. 25; International Covenant of Civil and Political Rights, Dec. 16, 1966, S. Exec. Rep. 102-23, 999 U.N.T.S. 171, art. 2.3 [hereinafter ICCPR]; ECHR *supra* note 3, Art. 13; International Convention on the Elimination of All Forms of Racial Discrimination, Dec. 21, 1965, S. Exec. Doc. C, 95-2 (1978), 660 U.N.T.S. 195, 212, art. 6.

⁶ U.N. HUMAN RIGHTS COMM., General Comment No. 29, States of Emergency (art. 4), U.N. Doc. CCPR/C/21/Rev.1/Add.11 ¶¶ 14-15 (Aug. 31, 2001).

clause allowing the HRC to request so. This paper also submits that, in terms of the substantive dimension of the right to an effective remedy, it is interesting to inquire about the converse influence. The jurisprudence of the IACtHR may have impacted on the remedial approach taken by the HRC. Alternatively, and in any case, it should serve as inspiration for the HRC for the development of the reparations.

Consequently, this paper considers the following questions: how has the right to an effective remedy become the normative source of other rights and obligations of a complex nature in the framework of the HRC case law? To what extent has this influenced the IACtHR's jurisprudence? What place does the right to an effective remedy hold in the jurisprudence of the HRC and the IACtHR? And what are the normative and practical consequences of extending the content and scope of such a right for both systems?

This article, therefore, adopts a comparative approach, focused on the HRC's views on individual communications and on the IACtHR's judgments. It is not the object of this paper to exhaustively analyze each of the rights and obligations identified as deriving from the right to an effective remedy, but rather to outline the development of the latter in the framework of both aforementioned systems. Therefore, the duty to carry out an investigation, to prosecute, to punish and to provide truth and reparation will not be addressed in an extensive manner, as each of these issues attract debates which exceed the limits of this paper. The present analysis will also concentrate on a specific type of human rights violation, namely, those violations that can be qualified as grave or gross. The specific focus is justified, given that the process of broadening the right to an effective remedy has been directly related to attempts to address those kinds of violations.

The discussion proceeds as follows: Part I provides clarification regarding the concept of remedies and the scope of the right to an effective remedy within the ICCPR. Part II examines the HRC's expansive interpretation of what can be considered, in principle, the procedural dimension of the right to an effective remedy. The duty to investigate, prosecute and punish, as well as the so-called right to truth, developed under the HRC's case law, are analyzed in this section. Part III addresses the influence of the HRC's interpretation upon the IACtHR's jurisprudence. In order to demonstrate this, the content and origin of the right to an effective remedy within the American Convention on Human Rights (ACHR) must first be explained. Subsequently, the duty to investigate, prosecute and punish, as well as the right to truth are again considered, this time within the framework of the IACtHR. Part IV discusses the substantive dimension of the right to an effective remedy, which has also been expanded by the HRC. This necessitates a consideration of the evolution of the requests for reparations which have progressively been made by the HRC, which in turn requires reference to be made to the IACtHR's jurisprudence. Accordingly, the possible influence of the IACtHR's case law on the HRC's practice on reparations is discussed. Finally, the last part of this paper proposes some conclusions, analysing the

main consequences which have flowed from the significant expansion of the right to an effective remedy and from the actual and potential mutual impacts of the HRC's and the IACtHR's jurisprudence.

II. REMEDIES AND THE RIGHT TO AN EFFECTIVE REMEDY UNDER THE ICCPR

The concept of remedy as a right has two dimensions: a procedural dimension and a substantive dimension.⁷ The first consists of ensuring individuals' access to independent and competent authorities that are capable of fairly deciding upon a claim of violation of their rights.⁸ The second, on the other hand, refers to the relief or redress afforded to a person who has been found to be a victim of a rights violation.⁹ In international human rights law, the latter notion of remedy is also known as reparation, and is normally the outcome of proceedings in which the State is found to be responsible for the violation of human rights.

The ACHR and the European Convention on Human Rights (ECHR) contain separate provisions for the right to an effective remedy (articles 25 and 13 respectively) and the right to reparation (articles 63.1 of the ACHR and 41 of the ECHR). However, the ICCPR does not reflect this distinction; article 2(3) of that Covenant embraces both meanings, according to the HRC's interpretation of that article.¹⁰ As a result, it will be seen that within the case law of the HRC it is difficult to delineate the procedural and the substantive dimensions of the right to an effective remedy. Very often both aspects appear to overlap and work together as a procedural tool for both the enforcement of rights and as a reparation measure. Article 2(3) of the ICCPR provides:

Each State Party to the present Covenant undertakes:

- (a) To ensure that any person whose rights or freedoms as herein recognized are violated shall have an effective remedy, notwithstanding that the violation has been committed by persons acting in an official capacity;
- (b) To ensure that any person claiming such a remedy shall have his right thereto determined by competent judicial, administrative or legislative authorities, or by any other competent authority provided for by the legal system of the State, and to develop the possibilities of judicial remedy;

⁷ SHELTON, *supra* note 4, at 7, 114; THOMAS ANTKOWIAK, *Remedial Approaches to Human Rights Violations: The Inter-American Court of Human Rights and Beyond*, in 46 COLUM. J. OF TRANSNAT'L L., 353, 356 (2008).

⁸ *Id.*

⁹ *Id.*

¹⁰ U.N. HUMAN RIGHTS COMM., General Comment No. 31, *The Nature of the General Legal Obligation Imposed on States Parties to the Covenant*, U.N. Doc. CCPR/C/21/Rev.1/Add.13 ¶¶ 15-6 (Mar. 29, 2004).

(c) To ensure that the competent authorities shall enforce such remedies when granted.

This provision is based on article 13 of the ECHR, although its scope is broader than the latter in paragraphs (b) and (c). In turn, article 2(3) of the ICCPR was incorporated into article 25 of the ACHR, though its introduction was not really consistent with the text of the ACHR, as will be explained below.¹¹ The wording of article 2(3) of the ICCPR as well as the *travaux préparatoires* of this provision indicate that the institutions entrusted with the power to declare whether a violation has taken place and to offer redress may be of a judicial, administrative or political nature.¹² These procedures involving ‘competent authorities’ have been understood broadly as encompassing different kinds of mechanisms, including administrative courts, inquiries by parliamentary commissions, inspectors and ombudsmen, informal preventive measures and judicial proceedings.¹³

The variety of possibilities for ensuring an adequate remedy is a consequence of the requirement of effectiveness. The appropriate form of procedural remedy may depend upon what will be ‘effective’ in the particular circumstances of the case. An effective remedy will be one which in practice brings the violation to an end and/or provides redress for a particular violation.

III. THE EXPANSION OF THE RIGHT TO A PROCEDURAL REMEDY IN THE HRC’S CASE LAW

Notwithstanding the above, judicial remedies are considered the ideal, as is evident from the explicit agreement between the States to “develop the possibilities of judicial remedy.” Moreover, the obligation to provide a judicial remedy is categorical when it comes to serious human rights violations such as executions, torture, enforced disappearances or human trafficking. In such cases, purely disciplinary and administrative remedies cannot replace judicial proceedings and cannot constitute an effective remedy.¹⁴

As a result of this interpretation, the right set out in article 2(3) of the ICCPR has undergone expansion. In fact, this provision has been

¹¹ See *infra* Section III.A.

¹² During the discussion of the draft, the U.K. proposal to establish a right to a judicial remedy was finally abandoned and instead decisions made by administrative and political organs were also accepted as effective remedies. Although this proposal- made by continental European and Latin American countries- prevailed, it was agreed to set forth a progressive obligation to develop judicial remedies. See MANFRED NOWAK, U.N. COVENANT ON CIVIL AND POLITICAL RIGHTS: CCPR COMMENTARY 32-4, 63-4 (2d ed. 2005).

¹³ *Id.* at 64-5.

¹⁴ U.N. HUMAN RIGHTS COMM., *Bautista de Arellano v. Colombia*, Communication No. 563/1993, U.N. Doc. CCPR/C/55/D/563/1993 ¶¶ 8.2, 10 (Oct. 27, 1995); General Comment No. 31, *supra* note 10, ¶ 18.

linked with a right of victims to access justice and more precisely, with the States' duty to investigate, prosecute and punish grave human rights violations as well as with the so-called right to truth. These two significant off-shoots of the right to an effective remedy are analyzed below.

A. *THE STATE DUTY TO INVESTIGATE, PROSECUTE AND PUNISH*

In cases of serious human rights violations, the HRC has derived from the right to an effective remedy a State obligation to conduct a prompt, thorough, impartial, and independent investigation in order to determine the factual circumstances of the violation, and to identify those deemed responsible.¹⁵ This duty has also been expressly recognized in specific human rights instruments in relation to grave violations.¹⁶

It is noteworthy to observe the way in which the HRC's case law has evolved in this respect. In a first stage, the HRC had already established the aforementioned obligation to investigate based on the duty to protect or ensure the right to life and the right not to be subjected to torture or inhuman treatment.¹⁷ Thus, it declared that such rights were violated by States which did not carry out a proper investigation of a killing or an enforced disappearance.¹⁸ This approach recalls article 2(1) general obligation "to ensure" the rights recognized in the ICCPR, but this provision was hardly invoked in the HRC's views. In some of the first cases, the HRC limited itself to finding a violation of the right to life or to integrity, and did not order the State to conduct an investigation in

¹⁵ See, *inter alia*, General Comment No. 31, *supra* note 10, ¶ 15; U.N. HUMAN RIGHTS COMM., Vicente, Villafañe et al. v. Colombia, Communication No. 612/1995, U.N. Doc. CCPR/C/60/D/612/1995 ¶ 8.8 (July 29, 1997); U.N. HUMAN RIGHTS COMM., Ali Bashasha v. The Libyan Arab Jamahiriya, Communication No. 1776/2008, U.N. Doc. CCPR/C/100/D/1776/2008 ¶ 9 (Oct. 20, 2010); U.N. HUMAN RIGHTS COMM., Boucherf v. Algeria, Communication No. 1196/2003, U.N. Doc. CCPR/C/86/D/1196/2003 ¶¶ 9.9, 11 (Mar. 30, 2006).

¹⁶ See, *e.g.*, Convention Against Torture and other Cruel, Inhuman or Degrading Treatment or Punishment, Dec. 10, 1984, 1465 U.N.T.S. 85, 113, art. 12; International Convention for the Protection of all Persons from Enforced Disappearance, G.A. res. A/61/177, Doc.A/61/488, C.N.737.2008, Dec. 20, 2006, art. 12.

¹⁷ U.N. HUMAN RIGHTS COMM., General Comment No. 6, *The Right to Life* (art. 6), UN Doc. 30/04/82, ¶ 4 (Apr. 30, 1982).

¹⁸ U.N. HUMAN RIGHTS COMM., Edgardo Santullo v. Uruguay, Communication No. 9/1977, U.N. Doc. CCPR/C/8/D/9/1977 ¶ 11 (Oct. 26, 1979); U.N. HUMAN RIGHTS COMM., Grille Motta v. Uruguay, Communication No. 11/1977, U.N. Doc. CCPR/C/10/D/11/1977., ¶ 16 (July 29, 1980) (referring to the violation of article 7 and 10); U.N. HUMAN RIGHTS COMM., Joaquín Herrera et al. v. Colombia, Communication No. 161/1983, U.N. Doc. CCPR/C/31/D/161/1983, ¶ 11 (Nov. 2, 1987): '[...] [T]he Committee disclose violations of the Covenant with respect to: Article 6, because the State party failed to take appropriate measures to prevent the disappearance and subsequent killings of Jose Herrera and Emma Rubio de Herrera and to investigate effectively the responsibility for their murders'. Similarly, U.N. HUMAN RIGHTS COMM., Basilio Atachahua v. Peru, Communication No. 540/1993, U.N. Doc. CCPR/C/56/D/540/1993 ¶ 8.3 (Mar. 25, 1996).

compliance with its obligation to provide an effective remedy under article 2(3).¹⁹ However, the HRC started to do the latter in subsequent cases, mainly by the end of the 1980s.²⁰ The possibility for this lay in the framework of the early decision of the HRC not to confine itself to just declare its findings in relation to a particular case, but to also require States to provide appropriate redress for the violations established. But the fulfilment of the obligation to investigate was not only required by the HRC as a measure of redress for the violation committed.²¹ Additionally and in parallel, the duty to investigate serious violations of human rights began to be addressed mainly under the scope of the right to an effective remedy, read in conjunction with the substantive right in question.²²

As a corollary of the above obligation to conduct investigations and to discover the facts and the perpetrators of grave violations, the right to an effective remedy was considered incompatible with amnesty laws.²³ If they exclude the possibility of initiating investigations into serious human rights violations and clarifying what happened to victims, States parties to the ICCPR must avoid or invalidate amnesties and analogous legal

¹⁹ See *Edgardo Santullo v. Uruguay*, *supra* note 18, at ¶ 13; *Grille Motta v. Uruguay*, *supra* note 18, at ¶ 18.

²⁰ See *Joaquín Herrera et al. v. Colombia*, *supra* note 18, at ¶ 12; *Basilio Atachahua v. Peru*, *supra* note 118, at ¶ 10.

²¹ See the last paragraphs of most of the cases cited in this paper. In the vast majority of them, the HRC uses a similar formula, such as “the Committee therefore urges the State party ... to conduct a thorough and effective investigation into the disappearance and death of [the victim].”

²² U.N. HUMAN RIGHTS COMM., General Comment No. 20, *Replaces general comment 7 concerning prohibition of torture and cruel treatment or punishment* (art. 7), UN Doc. 10/03/92, ¶ 14 (Mar. 10, 1992): ‘Article 7 should be read in conjunction with article 2, paragraph 3, of the Covenant. [...] The right to lodge complaints against maltreatment prohibited by article 7 must be recognized in the domestic law. Complaints must be investigated promptly and impartially by competent authorities so as to make the remedy effective [...]’; U.N. HUMAN RIGHTS COMM., *Sathasivam v. Sri Lanka*, Communication No. 1436/2005, U.N. Doc. CCPR/C/93/D/1436/2005 ¶ 6.4 (July 8, 2008); U.N. HUMAN RIGHTS COMM., *Amirov v. Russian Federation*, Communication No. 1447/2006, U.N. Doc. CCPR/C/95/D/1447/2006 ¶ 11.4 (Apr. 2, 2009); U.N. HUMAN RIGHTS COMM., *Pestaño v. The Philippines*, Communication No. 1619/2007, U.N. Doc. CCPR/C/98/D/1619/2007 ¶ 7.6 (Mar. 23, 2010).

²³ U.N. HUMAN RIGHTS COMM., *Concluding observations on Algeria*, CCPR/C/DZA/CO/3 ¶ 7 (Dec. 12, 2007); U.N. HUMAN RIGHTS COMM., *Hugo Rodríguez v. Uruguay*, Communication No. 322/1988, U.N. Doc. CCPR/C/51/D/322/1988 ¶ 12.4 (July 19, 1994). See also U.N. ECON. & SOC. COUNCIL (ECOSOC), Commission on Human Rights, *Question of the impunity of perpetrators of human rights violations (civil and political)*, Revised final report prepared by Mr. Joinet pursuant to Sub-Commission decision 1996/119, U.N. Doc. E/CN.4/Sub.2/1997/20/Rev.1 ¶ 27 (Oct. 2, 1997).

arrangements. This position has been upheld in the HRC's most recent views.²⁴

Nevertheless, the scope of the right to an effective remedy does not end there. Since the case of *Bautista v. Colombia*, the HRC has argued that article 2(3) of the ICCPR gives rise to a State obligation to prosecute and punish the perpetrators of serious violations of human rights.²⁵ The adoption of this position is not insignificant if one takes into account that at the time of drafting, a proposal to expressly recognize criminal prosecution as an example of an effective remedy was not approved.²⁶ Further, one could ask whether prosecution and punishment is necessary for the enforcement of the right, or is a necessary means of redress. The HRC seems to have answered both questions in the positive. It has affirmed that the mechanism providing for the substantiation of alleged gross violations should be a judicial criminal prosecution, while it has also determined that this step should be taken as a way to provide relief. Thereby, in the opinion of the HRC, the duty to investigate, prosecute and punish is grounded either in article 2(3) taken together with a substantive right, or in article 2(3)(a) alone. The first line of reasoning is used by the HRC when it addresses those duties as part of the procedural dimension of the right to an effective remedy in conjunction with the right to life or the prohibition on torture and ill treatment. The second line of argument implies that article 2(3)(a) alone would authorize the HRC to request investigation, prosecution and punishment as a means of reparation.²⁷

The State obligation to investigate and prosecute, however, does not have a correlative individual right. The HRC has clarified that "the Covenant does not provide a right for an individual to require that the State

²⁴ U.N. HUMAN RIGHTS COMM., *Khirani v. Algeria*, Communication No. 1905/2009, U.N. Doc. CCPR/C/104/D/1905/2009, ¶ 7.10 (Mar. 26, 2012), (referring to the Ordinance No. 06-01 implementing the Charter for Peace and National Reconciliation).

²⁵ See *Bautista de Arellano v. Colombia supra* note 14, ¶ 8.6.

²⁶ The proposal suggesting the introduction of a new paragraph saying 'violators shall be swiftly brought to the law, especially when they are public officials' was rejected by 6 votes to 3, with 4 abstentions. See MARC BOSSUYT, GUIDE TO THE 'TRAVAUX PREPARATOIRES' OF THE INTERNATIONAL COVENANT ON CIVIL AND POLITICAL RIGHTS 65 (1987).

²⁷ This issue will be further analyzed in Part IV. See U.N. HUMAN RIGHTS COMM., *Krasovskaya v. Belarus*, Communication No. 1820/2008, U.N. Doc. CCPR/C/104/D/1820/2008 ¶ 8.3 (Mar. 26, 2012); U.N. HUMAN RIGHTS COMM., *Bousroual v. Algeria*, Communication No. 992/2001, U.N. Doc. CCPR/C/86/D/992/2001 ¶ 11 (Mar. 30, 2006); U.N. HUMAN RIGHTS COMM., *Zyuskin v. Russian Federation*, Communication No. 1605/2007, U.N. Doc. CCPR/C/102/D/1605/2007 ¶ 13 (July 19, 2011); U.N. HUMAN RIGHTS COMM., *Zhumbaeva v. Kyrgyzstan*, Communication No. 1756/2008, U.N. Doc. CCPR/C/102/D/1756/2008 ¶¶ 8.10, 10 (July 19, 2011).

party criminally prosecute another person.”²⁸ But at the same time, the HRC has stressed that it “nevertheless considers the State party duty-bound not only to conduct thorough investigations into alleged violations of human rights, particularly enforced disappearances and violations of the right to life, but also to prosecute, try and punish the culprits. Thus, the State party is also under an obligation to prosecute, try and punish those held responsible for such violations.”²⁹

Certainly, the simultaneous acceptance of both arguments might be difficult to reconcile, although it is not impossible. In this regard, one has to bear in mind that the Covenant in fact includes State obligations that do not necessarily match with a corresponding individual right, such as the duty to submit reports.³⁰ Yet, the HRC decision to found the State duty (without a correlative right) to prosecute and punish on the right to an effective remedy raises questions of consistency, insofar as the latter is indeed an individual right allowing persons to complain before a competent body and to be granted reparation. In that sense, a more persuasive basis for the duty to prosecute and punish could be found in the general obligations contained in article 2 paragraphs 1 and 2: the obligations to ensure and to adopt measures to give effect to the rights set forth in the ICCPR.³¹ However, this has not been the position adopted by the HRC.

B. THE RIGHT TO TRUTH

There is yet another aspect of the right to an effective remedy that shows how far the interpretation of such a right has developed. This is the recognition of the so-called right to truth. Originally rooted in international humanitarian laws,³² this right has increasingly evolved through the activities of different international human rights bodies.³³ Despite the fact that this right has not been named and identified as such by the HRC, it is submitted that the decision in the case of *Almeida de Quinteros v. Uruguay*, concerning an enforced disappearance, constitutes the

²⁸ U.N. HUMAN RIGHTS COMM., *R. A. V. N. et al. v. Argentina*, Communication No. 343/1988, U.N. Doc. CCPR/C/38/D/343/1988 ¶ 5.5 (Mar. 26, 1990); *Vicente, Villafañe et al. v. Colombia*, *supra* note 15, ¶ 8.8; U.N. HUMAN RIGHTS COMM., *Messaouda Grioua v. Algeria*, Communication No. 1327/2004, U.N. Doc. CCPR/C/90/D/1327/2004 ¶ 9 (July 10, 2007).

²⁹ *Messaouda Grioua v. Algeria*, *supra* note 28, at ¶ 9.

³⁰ See ICCPR *supra* note 5, art. 40.

³¹ In the same vein, ANJA SEIBERT-FORTH, *The Fight against Impunity under the International Covenant on Civil and Political Rights*, 6 MAX PLANCK YEARBOOK OF UNITED NATIONS LAW, 323-4 (2002).

³² See Protocol I to the Geneva Conventions *supra* note 3, art. 32; U.N. ECON. & SOC. COUNCIL [ECOSOC], Commission on Human Rights, *Study on the Right to the Truth*, Report of the Office of the United Nations High Commissioner for Human Rights, U.N. Doc. E/CN.4/2006/91 ¶ 6 (Feb. 8, 2006), (hereinafter, *Study on the Right to the Truth*).

³³ See INTERNATIONAL COMMISSION OF JURISTS, *supra* note 1, at 81-89.

HRC's first precedent acknowledging a right of victims of gross violations to know the truth.³⁴

In that case the HRC held that “[t]he Committee understands the anguish and stress caused to the mother by the disappearance of her daughter and by the continuing uncertainty concerning her fate and whereabouts. *The author has the right to know what has happened to her daughter.*”³⁵ As a result, the HRC required Uruguay to “establish what has happened to Elena Quinteros.” In that case, the State’s failure to provide information about the fate and whereabouts of the victim was addressed as a form of inhuman and cruel treatment suffered by her mother; an approach maintained in subsequent cases.³⁶

Likewise, the right to know the facts surrounding an enforced disappearance or an extra-judicial execution has also been considered by the HRC to be part of the effective remedies required to be provided by the responsible State. On the one hand, the HRC has referred to the States’ duty to establish the facts of a grave violation as well as to identify its perpetrators in the context of the analysis of the obligation to investigate grave violations and to combat impunity.³⁷ On the other hand, after declaring a breach of those obligations, the HRC usually demands from States the adoption of measures aimed at providing information about the violation, the burial site or location of the remains of a disappeared victim.³⁸ These actions are usually required from States as part of their obligation to provide a substantive effective remedy, as will be illustrated when discussing the practice of the HRC on reparations below. In this regard, the verification of the facts and full and public disclosure of the truth have been characterised as a form of redress, namely, satisfaction.³⁹

Even though the HRC has not explicitly acknowledged the existence of an autonomous right to truth under the ICCPR, its jurisprudence on this matter has served as a fundamental basis for the further conceptualization of this right that has taken place in international human rights

³⁴ *Id.*, at 84.

³⁵ U.N. HUMAN RIGHTS COMM., *Almeida de Quinteros v. Uruguay*, Communication No. 107/1981, U.N. Doc. CCPR/C/19/D/107/1981 ¶ 14 (July 21, 1983), (emphasis added).

³⁶ U.N. HUMAN RIGHTS COMM., *Aliboev v. Tajikistan*, Communication No. 985/2001, U.N. Doc. CCPR/C/85/D/985/2001 ¶ 6.7 (Oct. 18, 2005); U.N. HUMAN RIGHTS COMM., *Staselovich v. Belarus*, Communication No. 887/1999, U.N. Doc. CCPR/C/77/D/887/1999 ¶ 9.2 (Apr. 3, 2003).

³⁷ *See, inter alia*, *Hugo Rodríguez v. Uruguay*, *supra* note 23, at ¶¶ 12.2-12.4; *Amirov v. Russian Federation*, *supra* note 22, at ¶¶ 11.3-11.4.

³⁸ *See Staselovich v. Belarus*, *supra* note 36, ¶ 11; U.N. HUMAN RIGHTS COMM., *Sarma v. Sri Lanka*, Communication No. 950/2000, U.N. Doc. CCPR/C/78/D/950/2000 ¶ 11 (July 16, 2003); U.N. HUMAN RIGHTS COMM., *Khalilova v. Tajikistan*, Communication No. 973/2001, U.N. Doc. CCPR/C/83/D/973/2001 ¶ 9 (Mar. 30, 2005).

³⁹ U.N. GAOR, *Basic Principles and Guidelines on the Right to a Remedy and Reparation for Victims of Gross Violations of International Human Rights Law and Serious Violations of International Humanitarian Law*, 3rd Comm., 60th Sess., U.N. Doc. A/RES/60/147 principle 22 (2005), [Hereinafter *Basic Principles*].

law.⁴⁰ Within the United Nations system the right to truth is deemed an inalienable and autonomous right linked to various other rights, including the right to an effective remedy.⁴¹ Here, and at the regional level, the right to truth is presented with a double character: as an individual and a collective right. The first involves the duty to provide specific information to the victims and/or relatives about the concrete circumstances of a gross violation, the place of burial or the fate of the victim, the causes of the victimization, the progress of any investigation and even the identity of the perpetrators.⁴² The collective dimension of the right to truth entails a State duty to disclose to the whole society information about the reasons why gross violations occurred, and the circumstances in which they took place.⁴³

It will be seen below that this development and, especially, the interpretation of the right to an effective remedy as a legal basis for a right to truth, has also influenced the Inter-American System. Over many years the Inter-American Commission of Human Rights worked to foster the recognition of a right to truth by the IACtHR, until the latter finally acceded. In this respect, it is important to highlight that one of the first findings of a violation of the right to know the truth by the Inter-American Commission was partially founded on a precedent set by the HRC. In fact, in the case of *Ellacuría v. El Salvador*,⁴⁴ the Inter-American Commission invoked the *Almeida Quinteros v. Uruguay* case in support of its position. The impact that the HRC's position on both this issue and the duty to investigate and prosecute has had on the IACtHR's jurisprudence is the focus of the following section.

IV. THE INFLUENCE OF THE HRC'S EXPANSIVE INTER- PRETATION OF THE RIGHT TO AN EFFECTIVE REMEDY ON THE JURISPRUDENCE OF THE INTER-AMERICAN COURT

A. THE RIGHT TO AN EFFECTIVE REMEDY UNDER THE ACHR

As previously stated, the ACHR also contains an effective remedy provision, which is different from that governing reparations. The right to an effective remedy is enshrined in article 25 of the ACHR, the word-

⁴⁰ ECOSOC, *Study on the Right to the Truth*, *supra* note 32, at ¶¶ 25-31; THOMAS ANTKOWIAK, *Truth as Right and Remedy in International Human Rights Experience*, 23 MICH. J. INT'L L., 992 (2002).

⁴¹ ECOSOC, *Study on the Right to the Truth*, *supra* note 32, at ¶¶ 55-6.

⁴² U.N. ECON. & SOC. COUNCIL [ECOSOC], Commission on Human Rights, *Report of the Independent Expert to Update the Set of Principles to Combat Impunity*, Diane Orentlicher, U.N. Doc. E/CN.4/2005/102/Add.1, principle 4 (Feb. 8, 2005).

⁴³ *Id.*, principle 2; YASMIN NAQVI, *The Right to Truth in International Law: Fact or Fiction?*, 88 INT'L REV. OF THE RED CROSS, 260 (2006).

⁴⁴ *Ellacuría v. El Salvador*, Case No. 10.488, Inter-Am. Comm'n H.R., Report No. 136/99, OEA/Ser.L/V/II.106 Doc. 3 rev., (1999).

ing of which was mostly copied from article 2(3) of the ICCPR. The provision reads:

Article 25. Right to Judicial Protection

1. Everyone has the right to simple and prompt recourse, or any other effective recourse, to a competent court or tribunal for protection against acts that violate his fundamental rights recognized by the constitution or laws of the state concerned or by this Convention, even though such violation may have been committed by persons acting in the course of their official duties.

2. The States Parties undertake:

- a. to ensure that any person claiming such remedy shall have his rights determined by the competent authority provided for by the legal system of the state;
- b. to develop the possibilities of judicial remedy; and
- c. to ensure that the competent authorities shall enforce such remedies when granted.

A first glance shows the similarity between this provision and article 2(3) of the ICCPR, but also an important difference: unlike the latter, the former establishes a right to a “judicial” remedy. Therefore, the ACHR affords stronger protection by limiting the possibilities of remedies than can be considered appropriate for addressing a violation of human rights to those of a judicial nature. However, how can this be reconciled with the wording of article 25(2)(b), which places an obligation on the States to develop the *possibilities* of a judicial remedy? As seen above, such a statement was understandable in the context of the discussion about the ICCPR where the agreement reached was to give States a wide margin as to what kind of remedies they should provide. The idea to insert article 2(3) of the ICCPR into the text of article 25 of the ACHR emerged when drafters realized that the American provision did not mention the rights in the Convention amongst those protected by an effective remedy.⁴⁵ But instead of just copying the entire provision from the ICCPR, or only incorporating a mention of the rights in the Convention, the drafters added the wording of article 2(3), when the right to a “judicial remedy” was already agreed and established.⁴⁶ The result was a provision which lacks internal coherence, although its purpose was clearly to reinforce the State obligation to give effect to the rights recognised domestically and in the ACHR.

⁴⁵ ORGANIZATION OF AMERICAN STATES [OAS], Conferencia Especializada Interamericana sobre Derechos Humanos. Actas y Documentos, OEA/Ser.K/XVI/1.2, 41, 55-6 (Nov. 7-22, 1969).

⁴⁶ *Id.*, at 261-63; CECILIA MEDINA, LA CONVENCION AMERICANA DE DERECHOS HUMANOS: VIDA, INTEGRIDAD PERSONAL, LIBERTAD PERSONAL, DEBIDO PROCESO Y RECURSO JUDICIAL 368 (2003).

In reality, article 25 of the ACHR was created to set out a judicial remedy of a particular nature: the so-called *amparo*, “which is a simple and prompt remedy designed for the protection of all of the rights recognized by the constitutions and laws of the States Parties and by the Convention”.⁴⁷ This is a procedural mechanism of Latin-American origin, initially designed to protect constitutional rights.⁴⁸ Therefore, unlike article 2(3) of the ICCPR, its American counterpart only enshrined this specific form of judicial remedy to protect the rights recognized in domestic law and the Convention. Of course, this does not mean that States do not have a duty to provide other kinds of remedies which may require a longer or more complex process but simply that such obligation would not be grounded in article 25.⁴⁹

The introduction of article 2(3) of the ICCPR into article 25 of the ACHR provides the first significant reason to consider the interpretation of the right to an effective remedy adopted by the HRC. The IACtHR has repeatedly referred to the HRC’s case law to support its reasoning and decisions on a variety of issues,⁵⁰ and the examination of the right to

⁴⁷ Habeas Corpus in Emergency Situations (Arts. 27(2), 25(1) and 7(6) American Convention on Human Rights), Advisory Opinion 8/87, Inter-Am. Ct. H.R. (ser. A) No. 8 ¶ 32 (Jan. 30, 1987); Judicial Guarantees in States of Emergency (Arts. 27(2), 25 and 8 American Convention on Human Rights), Advisory Opinion 9/87, Inter-Am. Ct. H.R. (ser. A) No. 9 ¶¶ 23-4 (Oct. 6, 1987).

⁴⁸ CARLOS AYALA, *Del Amparo Constitucional al Amparo interamericano como Institutos para la Protección de los Derechos Humanos*, 2 LIBER AAMICORUM HECTOR FIX-ZAMUDIO 342-44, (Inter-Am. Ct. H.R. ed. 1998).

⁴⁹ The admissibility rule of exhaustion of domestic remedies and the right to fair trial reflect that States are obligated to provide an ample spectrum of mechanisms for bringing complaints of alleged violations of human rights; MEDINA, *supra* note 46, at 359-360.

⁵⁰ The IACtHR has expressly referred to the HRC case law in, *inter alia*, the following judgments: Durand and Ugarte v. Peru, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 68, ¶ 124 (Aug. 16, 2000); Cantoral-Benavides v. Peru, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 69, ¶¶ 86-102 (Aug. 18, 2000); Bámaca-Velásquez v. Guatemala, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 70, ¶¶ 153, 164 & 172 (Nov. 25, 2000); Herrera-Ulloa v. Costa Rica, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 107, ¶¶ 114, 166 (July 2, 2004); Ricardo Canese v. Paraguay, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 111, ¶¶ 84, 115-16, 124, 132 (Aug. 31, 2004); Tibi v. Ecuador, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 114, ¶ 186 (Sept. 7, 2004); De la Cruz-Flores v. Peru, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 115, ¶ 100 (Nov. 18, 2004); Lori Berenson-Mejia v. Peru, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 119, ¶ 154 (Nov. 25, 2004); Caesar v. Trinidad and Tobago, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 123, ¶ 62-3 (Mar. 11, 2005); Saramaka People v. Suriname, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 172, ¶¶ 94, 130 (Nov. 28, 2007); Radilla-Pacheco v. Mexico, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 209, ¶¶ 303-04 (Nov. 23, 2009); Gelman v. Uruguay,

an effective remedy has been no exception. Accordingly, it is important to be aware of the limitations of the approaches adopted by both bodies, and of the different scopes of article 25 of the ACHR and article 2(3) of the ICCPR. These factors help explain some of the problems that have arisen in the jurisprudence of the IACtHR as a result of trying to accommodate a similar notion of effective remedy to that developed in the HRC's case law.

B. *THE STATE DUTY TO INVESTIGATE, PROSECUTE AND PUNISH*

The initial consideration by the IACtHR of a State duty to seriously investigate grave violations of human rights took place in its landmark case of *Velásquez Rodríguez*.⁵¹ However, the legal justification for such an obligation was not built upon the right to an effective remedy, but on the obligation to guarantee or ensure the rights provided for in article 1(1) of the ACHR. As in the first cases decided by the HRC, the IACtHR found that the obligation to protect the right to life and to personal integrity required States to initiate diligent and prompt investigations if it was alleged that those rights had been violated.⁵² Moreover, in that case, the Court derived from article 1(1) not only a duty to investigate, but also to prosecute and punish those found responsible for the enforced disappearance of Manfredo Velásquez.⁵³ Consequently, the Court concluded that the lack of an adequate and thorough investigation, prosecution and punishment constituted a breach of the substantive right to life, read together with article 1(1).⁵⁴

In later cases, the Court rejected the arguments of the Inter-American Commission to develop a duty to investigate, prosecute and punish on the basis of article 25, reaffirming that this provision concerned a simple and prompt recourse such as an *habeas corpus*.⁵⁵ However, in some of these cases the IACtHR founded the obligation to inves-

Merits and Reparations Judgment, Inter-Am. Ct. H.R., (ser. C) No. 221, ¶¶ 205-07 (Feb. 24, 2011).

⁵¹ *Velásquez-Rodríguez v. Honduras*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 4 (July 29, 1988). The same reasoning was applied in *Godínez Cruz v. Honduras*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 5, ¶¶ 187-97 (Jan. 20, 1989).

⁵² *Id.* at ¶¶ 166, 174, 177.

⁵³ *Id.* at ¶¶ 174, 180, 188.

⁵⁴ *Id.* at ¶¶ 188, 194.4.

⁵⁵ See *Genie Lacayo v. Nicaragua*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 30, ¶ 89 (Jan. 29, 1997); *Blake v. Guatemala*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 36, ¶¶ 91, 98, 104 (Jan. 24, 1998); *The "White Van" (Paniagua Morales et al.) v. Guatemala*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 37, ¶¶ 159-60, 164-68 (Mar. 8, 1998). In other cases the ACtHR only applied art. 25 to assess the availability and/or effectiveness of habeas corpus recourses: *Loayza Tamayo v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 33, ¶¶ 52-55 (Sept. 17, 1997); *Castillo Páez v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 34, ¶¶ 82-84 (Nov. 3, 1997); *Castillo Petruzzi et al. v. Peru*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 52, ¶¶ 180-88 (May 30, 1999).

tigate, prosecute and punish grave violations of human rights in article 8(1) (right to a fair trial) and not in the obligation to ensure the rights to life or to personal integrity. This interpretation was held in two cases where the IACtHR determined that it lacked jurisdiction to rule on the legal implications of the main victims' death, given that their murder took place prior to the State acceptance of the IACtHR's jurisdiction.⁵⁶ As a consequence, the Court did not address the obligation to ensure the right to life of the deceased victims but instead focused on their relatives' right to a fair hearing. Interestingly, this argumentation was further extended to other cases where such a jurisdictional problem was not at stake and subsequently was complemented with the right to an effective remedy. Thus, the Court finally shifted its position and founded the obligation to investigate, prosecute and punish in the right to a fair hearing and the right to an effective remedy taken together.

In fact, in the case of *Durand and Ugarte v. Peru*,⁵⁷ the IACtHR addressed the deficiencies of the criminal investigation and prosecution in that case under both the right to judicial protection (or effective remedy) and the right to a fair trial (article 8). In the opinion of the Court, "article 8(1) of the American Convention, in connection with Article 25(1) thereof, confers to victims' relatives the right to investigate their disappearance and death by State authorities, to carry out a process against the liable parties of unlawful acts, to impose the corresponding sanctions, and to compensate damages suffered by their relatives."⁵⁸ It is interesting to note that in support of this new approach, the Court referred to the case law of the HRC, according to which States parties have a duty to investigate, to prosecute and to punish those deemed responsible for grave violations, because "purely disciplinary and administrative remedies cannot be deemed to constitute adequate and effective remedies within the meaning of article 2, paragraph 3 [of the ICCPR]."⁵⁹ However, it should be pointed out as well that the IACtHR has gone further than the HRC, because unlike the latter, the former has recognized that

⁵⁶ *Genie Lacayo v. Nicaragua*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 30, ¶ 89 (Jan. 29, 1997). See also *Blake v. Guatemala*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 36, ¶¶ 91, 98, 104 (Jan. 24, 1998). A similar but less clear suggestion can be found in *The "White Van" (Paniagua Morales et.al) v. Guatemala*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 37, ¶¶ 139, 142-42, 155-56 (Mar. 8, 1998).

⁵⁷ *Durand and Ugarte v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 68, ¶ 124 (Aug. 16, 2000).

⁵⁸ *Id.* at ¶¶ 130-31.

⁵⁹ In ¶ 124 of the *Durand and Ugarte* case, the IACtHR referred to the following views of the HRC: *Vicente, Villafañe et al. v. Colombia*, Communication No. 612/1995, U.N. Doc. CCPR/C/60/D/612/1995 ¶ 8.8 (July 29, 1997); *Bautista de Arellano v. Colombia*, Communication No. 563/1993, U.N. Doc. CCPR/C/55/D/563/1993 ¶¶ 8.2, 10 (Oct. 27, 1995).

there is a right of victims to require a criminal investigation and prosecution of those involved in the violation from States.⁶⁰

It is clear from the *Durand and Ugarte v. Peru* judgment that the IACtHR understood the right to an effective remedy under the ACHR as embracing much more than simply the *amparo* recourse, as it was originally conceived. The Court seemed to identify article 25 of the ACHR as the legal source of a range of recourses which States should provide, including criminal proceedings.⁶¹ This led the Court to declare that a “right to access to justice” was contained in the right to judicial protection contained in article 25.⁶² Therefore, the Court departed from the primary notion of article 25 as a “simple and prompt recourse” and moved closer to the understanding of an effective remedy elaborated by the HRC within the framework of the ICCPR.

The position taken in the *Durand and Ugarte* case has been maintained in subsequent judgments of the IACtHR, although in some cases concerning murders or executions the IACtHR has acknowledged the obligation to guarantee the right to life as the main source of the State obligation to investigate, prosecute and punish.⁶³ Still in most cases, particularly when it comes to the rights of the victims’ next of kin, the investigation, prosecution and punishment of grave violations of human rights are analyzed under a section devoted to article 8 and 25, taken together. This combined assessment of the obligations under those rights makes it very difficult to distinguish the scope and the role of each. The IACtHR early recognized that article 25 and 8 are interconnected, as they are too with article 1(1).⁶⁴ A precondition of any effective remedy is its substantiation according to the due process of law.⁶⁵ However, the IACtHR’s jurisprudence has made of these rights a “conceptually organic whole.”⁶⁶

⁶⁰ See, *inter alia*, *Durand and Ugarte v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 68, ¶ 130 (Aug. 16, 2000); *The 19 Tradesmen v. Colombia*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 109, ¶ 187 (July 5, 2004); *Radilla-Pacheco v. Mexico*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 209, ¶180 (Nov. 23, 2009).

⁶¹ This understanding was subsequently more explicitly confirmed in *The Moiwana Community v. Surinam*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 124, ¶ 148 (June 15, 2005).

⁶² *Cantos v. Argentina*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 97, ¶ 52 (Nov. 28, 2002).

⁶³ See, *inter alia*, *Kawas-Fernández v. Honduras*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 196, ¶¶ 72-108 (April 3, 2009).

⁶⁴ *Judicial Guarantees in States of Emergency* (Arts. 27(2), 25 and 8 American Convention on Human Rights), Advisory Opinion 9/87, Inter-Am. Ct. H.R. (ser. A) No. 9 ¶¶ 23-4, 27-8 (Oct. 6, 1987).

⁶⁵ *Id.* at ¶¶24, 27.

⁶⁶ ANTONIO CANÇADO, *THE ACCESS OF INDIVIDUALS TO INTERNATIONAL JUSTICE* 66 (2011). This former president of the IACtHR is one of the main proponents of a joint analysis of arts. 25 and 8 of the ACHR. He also advanced the right to access to justice as a right grounded in both provisions.

Leaving aside the controversial issue of a right of victims to State prosecution and punishment, in the current reasoning of the Court the limits of article 25 and 8 of the ACHR seem to be blurred.⁶⁷ Moreover, this suggests that such a victim right can be based on two procedural provisions that basically provide guarantees for the enforcement of rights but give scarce indication about the substantive content of any right.⁶⁸ From another point of view, the stronger condemnation which flows from a finding by the HRC of a violation of a substantive right (such as to life or integrity), read together with article 2(3) of the ICCPR, as compared to that which flows from a determination there has been a violation of article 25 and 8 of the ACHR, cannot be overlooked.

In line with the assertion of the duties to investigate, prosecute and punish grave violations of human rights, the IACtHR has also affirmed the incompatibility of amnesty laws with the ACHR. Since the Court adopted its broad understanding of article 25 and its examination of that article in conjunction with article 8, the proscription of amnesties -similarly to the approach taken by the HRC- has been justified based on the right to an effective remedy, though in conjunction with the right to a fair trial and in light of articles 1(1) and 2 of the ACHR.⁶⁹

A 2010 judgment clearly demonstrates this trend. In the case of *Gomes-Lund v. Brazil*,⁷⁰ in the section of the judgment addressing article 25 together with article 8, the Court stated:

The obligation to investigate, and where applicable, punish the serious violations of human rights have been affirmed by all of the international systems for the protection of human rights. In the universal system, the United Nations Human Rights Committee [...] considered in its constant jurisprudence that the criminal investigation and the ensuing prosecution are corrective measures that are necessary for violations of human rights [...] [and] concluded that States must establish what has oc-

⁶⁷ The former president of the IACtHR, Cecilia Medina, is very critical of this unified treatment. See MEDINA, *supra* note 46, at 364-65, 371, 380; *The 19 Tradesmen v. Colombia*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 109, ¶ 187 (July 5, 2004). (Medina, J., dissenting).

⁶⁸ *The 19 Tradesmen v. Colombia*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 109, ¶ 7 (July 5, 2004). (Medina, J., dissenting).

⁶⁹ See, e.g. *Case of Barrios Altos v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 75, ¶ 42-3 (March 14, 2001); *La Cantuta v. Peru*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 162, ¶ 226 (Nov. 29, 2006); *Chitay Nech et al. v. Guatemala*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 212, ¶ 199 (May 25, 2010); *Radilla-Pacheco v. Mexico*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 209, ¶ 212 (Nov. 23, 2009).

⁷⁰ *Gomes-Lund et al. (Guerrilha do Araguaia) v. Brazil*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 219 ¶ 126 (Nov. 24, 2010).

curred to the disappeared victims and bring justice to those responsible.⁷¹

A few paragraphs later, the Court continues: “Likewise, [...] in the case of *Hugo Rodríguez v. Uruguay*, it [the HRC] noted that it cannot accept the posture of a State of not being obligated to investigate human rights violations committed during a prior regime given an amnesty law, and it reaffirmed that amnesties in regards to serious human rights violations are incompatible with the International Covenant on Civil and Political Rights”.⁷² The Court concludes by adopting the same position as the HRC in that case, and declaring a violation of articles 8 and 25 of the ACHR because the Brazilian amnesty law breached the State obligation to investigate and punish.⁷³ The same arguments were reiterated in the case of *Gelman v. Uruguay*.⁷⁴

C. THE RIGHT TO TRUTH

Although the IACtHR initially rejected the Inter-American Commission’s position that the ACHR protected a right to truth,⁷⁵ it finally recognized such a right through its joint interpretation of the rights to an effective remedy and to a fair trial.⁷⁶ Thus, the other side of the State duty to thoroughly investigate grave violations of human rights was a right to truth. Notwithstanding that the Inter-American Commission’s arguments related to the right to truth took into account the case law of the HRC, the IACtHR’s recognition of the right went much further than the HRC’s interpretation. Unlike the HRC, the IACtHR has explicitly acknowledged a right of victims of serious violations (and their relatives) to know the truth and, moreover, it has developed this concept in remarkable detail. According to the IACtHR, the right to truth is not an autonomous right in the ACHR, but it is subsumed within the right to access to justice contained in articles 25 and 8, read together, which allows victims and relatives to require the State to investigate and prosecute.⁷⁷ The right to truth is, in the

⁷¹ *Id.* at ¶ 141.

⁷² *Id.* at ¶ 157.

⁷³ *Id.* at ¶¶ 172, 180.

⁷⁴ *Gelman v. Uruguay*, Merits and Reparations Judgment, Inter-Am. Ct. H.R., (ser. C) No. 221, ¶¶ 205-6, 227 (Feb. 24, 2011).

⁷⁵ *Castillo-Páez v. Peru*, Merits Judgment, Inter-Am. Ct. H.R., (ser. C) No. 34, ¶ 85-6 (Nov. 3, 1997).

⁷⁶ *Mack-Chang v. Guatemala*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 101, ¶¶ 273-74 (Nov. 25, 2003); *Gomes-Lund et al. (Guerrilha do Araguaia) v. Brazil*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 219, ¶¶ 200-1 (Nov. 24, 2010); *Gelman v. Uruguay*, Merits and Reparations Judgment, Inter-Am. Ct. H.R., (ser. C) No. 221, ¶¶ 243-4 (Feb. 24, 2011).

⁷⁷ *The Pueblo Bello Massacre v. Colombia*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 140, ¶¶ 219-20 (Jan. 31, 2006); *Baldeón-García v. Peru*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 147, ¶ 166 (Apr. 6, 2006); *Montero-Aranguren et al. (Detention Center of Catia) v. Venezuela*,

opinion of the Court, also related to the State obligation to ensure the rights and repair their violation, as well as the right to seek and receive information.⁷⁸ Further, it is conceptualized as a twofold right. The right to truth encompasses an individual and a collective right belonging to the victims and their relatives on the one hand, and to society as a whole on the other.⁷⁹ Both elements of the right are realized through conducting diligent criminal investigations and by public dissemination of the results obtained from that and other investigative procedures.

By merging article 25 and 8 of the ACHR, the IACtHR has combined multiple, interconnecting rights and duties –the duty to investigate, prosecute, punish and repair and the right to truth under the umbrella of “access to justice”. A 2011 judgment concerning an enforced disappearance clearly demonstrates this type of reasoning. In the case of *Contreras et al. v. El Salvador*, the Court explained that “[...] the long periods of procedural inactivity, the refusal to provide information on the military operations, and the lack of diligence and exhaustiveness in the investigations by the authorities in charge of them, permit the Court to conclude that all the domestic proceedings have not constituted effective remedies to determine the fate or to discover the whereabouts of the victims, or to guarantee the rights of access to justice and to know the truth, through the investigation and eventual punishment of those responsible, and full reparation of the consequences of the violations.”⁸⁰

The arguments developed by the IACtHR in this matter illustrate a broad interpretation of article 25, close in many aspects to the evolution of the right to an effective remedy within the HRC’s case law. Nevertheless, it should be recalled that in the latter’s context, the wide meaning of the right to an effective remedy is the result of various factors. Among them, 1) the

Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 150, ¶ 55 (July 5, 2006); *Servellón-García et al. v. Honduras*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 152, ¶ 76 (Sept. 21, 2006).

⁷⁸ In *Gomes-Lund et al. (Guerrilha do Araguaia) v. Brazil*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 219, ¶ 201 (Nov. 24, 2010), the Court noted that because of the facts involved, the right to know the truth was related to an action brought by relatives to access certain information, as well as related to access to justice and the right to seek and receive information as enshrined in art. 13 of the Convention. Accordingly, the Court analyzed the right to truth under art. 13 ‘Freedom of Thought and Expression.’

⁷⁹ See, e.g., *Gomes-Lund et al. (Guerrilha do Araguaia) v. Brazil*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 219, ¶200 (Nov. 24, 2010); *Mack-Chang v. Guatemala*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 101, ¶¶ 276 (Nov. 25, 2003); *Gelman v. Uruguay*, Merits and Reparations Judgment, Inter-Am. Ct. H.R., (ser. C) No. 221, ¶ 273 (Feb. 24, 2011); *Contreras et al. v. El Salvador*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 232, ¶ 170 (Aug. 31, 2011).

⁸⁰ *Contreras et al. v. El Salvador*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 232, ¶ 176 (Aug. 31, 2011).

acceptance of an ample range of procedural avenues as effective remedies in terms of article 2(3) of the ICCPR; 2) the need for the HRC to ground its reparation requests in that same provision and 3) the lack of jurisprudential development of the obligation to ensure the rights in article 2 (1) of the ICCPR.

In its jurisprudence the IACtHR continues to mention article 1(1) –the obligation to ensure the rights – in addition to articles 25 and 8 in support of a right to require investigation, prosecution and punishment as well as a right to truth. This development has brought with it definition, with remarkable detail, of the States’ obligation to investigate.⁸¹ It has also underscored that States are bound to ensure much more than formal access to their justice system.⁸² However, the IACtHR’s joint approach to these three provisions offers an overlapping interpretation of the content of each that raises some problematic issues. First, that approach tends to merge the legal standards applicable to the due process of law and those governing the right to an effective remedy. This could have not only theoretical/technical consequences, but also practical.⁸³ Second, if this trend continues, one may ask whether the possibility exists that the obligation to ensure in article 1 could progressively lose the meaningful content that it has had.

V. THE EXPANSION OF THE SUBSTANTIVE NOTION OF A REMEDY: REPARATIONS IN THE HRC AND THE POSSIBLE IMPACT OF THE IACTHR’S JURISPRUDENCE

As indicated at the outset of this analysis the substantive component of the right to an effective remedy consists in providing redress when a human rights violation has been established. This form of remedy has a crucial place in international law, in that it reflects the general principle of international law that “any breach of an engagement involves an obligation to make reparation.”⁸⁴ The reparation has to be provided with a view to rectifying the consequences of the wrongdoing, and in order to re-establish the situation that existed prior to the commission of the illegal act. This is what is known as *restitution in integrum* or the “full remedy

⁸¹ Especially since *Juan Humberto Sánchez v. Honduras*, Preliminary Objection, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 99, ¶¶ 127-28 (June 7, 2003).

⁸² CANÇADO, *supra* note 66 at 71.

⁸³ Medina, *supra* note 46, at 365, 371. An example provided by her is very illustrative. She suggests that if the “reasonable time” criterion from article 8 is applied to the judicial remedy in article 25, the time of substantiation of the latter will be measured against the duration of civil and criminal procedure which, by nature, are far from being as “prompt and simple” as the remedy in article 25 was meant to be.

⁸⁴ *Factory at Chorzow (Germany v. Poland)*, P.C.I.J. (ser. A) No. 17, at 29, (Sept. 13, 1928); *Gomes-Lund et al. (Guerrilha do Araguaia) v. Brazil*, Preliminary Objections, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 219, ¶ 245, n 369 (Nov. 24, 2010).

rule”,⁸⁵ which is also applicable in the context of violations of human rights obligations. However, in this latter sphere, such realization is extremely difficult to achieve.

A. REPARATIONS AT THE HRC

The best example of a broad reading of the right to an effective remedy so as to include the State obligation to provide and the corresponding right of the victim to receive reparation is contained in the practice of the HRC. The reason for this is simple: the ICCPR, unlike the ACHR and the ECHR, does not contain a provision concerning reparation for violations of the rights set forth in that covenant. There is no equivalent to article 63 of the ACHR or article 41 of the ECHR in the ICCPR or its First Optional Protocol (hereinafter OP). This asymmetry is probably due to the different nature of the HRC, as compared to the IACtHR and the European Court of Human Rights. The HRC is not a tribunal and, therefore, it does not issue binding judgments as regional human rights courts do, and nor, in principle, would it have the power to order reparations.⁸⁶ However, the HRC clearly performs an adjudicative function when considering individual communications, which as such requires a pronouncement on the remedies to be afforded to victims.⁸⁷ Moreover, the HRC has gone through a process of increasing judicialization, whereby its decisions have become very similar to binding judgments in form and substance.⁸⁸

Thus, despite the silence of the ICCPR and its OP, the HRC has interpreted article 2(3), the right to an effective remedy, as the normative source for requesting States parties to repair the violations established in its views on individual communications.⁸⁹ The HRC adopted this position very early in its practice,⁹⁰ although it was not self-evident that the HRC’s mandate could go beyond finding violations and extend to recommending reparations.⁹¹ However, the position taken by the HRC on this issue must be

⁸⁵ SONJA STARR, *Rethinking ‘Effective Remedies’: Remedial Deterrence in International Courts*, 83 N.Y.U. L. REV. 699 (2008).

⁸⁶ NOWAK, *supra* note 4, at 75.

⁸⁷ THOMAS BUERGENTHAL, *The U.N. Human Rights Committee*, 5 MAX PLANCK YEARBOOK OF UNITED NATIONS LAW 396 (2001); SEIBERT-FORTH, *supra* note 31, at 309.

⁸⁸ YOGESH TYAGI, *THE U.N. HUMAN RIGHTS COMMITTEE. PRACTICE AND PROCEDURE* 562 (1st ed. 2011).

⁸⁹ While at the beginning it was not specified on which grounds the HRC was requiring States to remedy the violations found, this was subsequently defined as art. 2(3) and later, more precisely, art. 2(3)(a). The latter is the current practice.

⁹⁰ See U.N. HUMAN RIGHTS COMM., *Bazzano/Valentini/Massera v. Uruguay*, Communication No. 5/1977, U.N. Doc. CCPR/C/7/D/5/1977, ¶ 10 in fine (Aug. 15, 1979).

⁹¹ In fact, it was even argued that such a possibility would be in violation of article 2(7) of the UN Charter, as it would constitute intervention in the domestic affairs of State parties. See JAKOB TH. MÖLLER AND ALFRED DE ZAYAS, *UNITED NATIONS HUMAN RIGHTS COMMITTEE. CASE LAW 1977-2008. A HANDBOOK* 456 (2009).

viewed in light of the type of violations that it examined during its first period of activity, that is, gross human rights violation committed in the context of dictatorships.⁹²

As a result of its interpretation of article 2(3), the HRC has required the implementation of different sorts of reparation measures. Although initially the HRC mainly focused on granting compensation (which is still the most common measure), it has progressively demanded other forms of redress, such as the adjustment of domestic legislation⁹³ and the release of a detainee.⁹⁴ The type of reparation demanded depends on the nature of the right involved and the features of the violation. To date a considerable variety of measures have been recommended by the HRC, and over time they have become more specific. Among the measures which have been required by the HRC as “effective remedies” for violations are: the nullification of a conviction and refund of a fine paid by the victim;⁹⁵ restraint from enforcement and revocation of an expulsion order;⁹⁶ a public apology;⁹⁷ commutation of a death sentence;⁹⁸ early consideration for parole;⁹⁹ retrial under due judicial guarantees;¹⁰⁰ protection from threats;¹⁰¹

⁹² Most of them against Uruguay.

⁹³ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Salgar de Montejo v. Colombia, Communication No. 64/1979, U.N. Doc. CCPR/C/15/D/64/1979, ¶ 12 (Mar. 24, 1982); U.N. HUMAN RIGHTS COMM., Karakurt v. Austria, Communication No. 965/2000, U.N. Doc. CCPR/C/74/D/965/2000, ¶ 10 (Apr. 4, 2002).

⁹⁴ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Polay Campos v. Peru, Communication No. 577/1994, U.N. Doc. CCPR/C/61/D/577/1994, ¶ 10 (Nov. 6, 1997); U.N. HUMAN RIGHTS COMM., Koreba v. Belarus, Communication No. 1390/2005, U.N. Doc. CCPR/C/100/D/1390/2005, ¶ 9 (Oct. 25, 2010).

⁹⁵ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Coleman v. Australia, Communication No. 1157/2003, U.N. Doc. CCPR/C/87/D/1157/2003, ¶ 9 (July 17, 2006); U.N. HUMAN RIGHTS COMM., Bodrozic v. Serbia and Montenegro, Communication No. 1180/2003, U.N. Doc. CCPR/C/85/D/1180/2003, ¶ 9 (Oct. 31, 2005).

⁹⁶ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Byahuranga v. Denmark, Communication No. 1222/2003, U.N. Doc. CCPR/C/82/D/1222/2003, ¶ 13 (Nov. 1, 2004); U.N. HUMAN RIGHTS COMM., Bakhtiyari v. Australia, Communication No. 1069/2002, U.N. Doc. CCPR/C/79/D/1069/2002, ¶ 11 (Oct. 29, 2003).

⁹⁷ U.N. HUMAN RIGHTS COMM., William Lecraft v. Spain, Communication No. 1493/2006, U.N. Doc. CCPR/C/96/D/1493/2006 ¶ 9 (July 27, 2009).

⁹⁸ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Lubuto v. Zambia, Communication No. 390/1990, U.N. Doc. CCPR/C/55/D/390/1990, ¶ 9 (Oct. 31, 1995); U.N. HUMAN RIGHTS COMM., Rayos v. Philippines, Communication No. 1167/2003, U.N. Doc. CCPR/C/81/D/1167/2003, ¶ 9 (July 27, 2004); U.N. HUMAN RIGHTS COMM., Chisanga v. Zambia, Communication No. 1132/2002, U.N. Doc. CCPR/C/85/D/1132/2002, ¶ 9 (Oct. 18, 2005).

⁹⁹ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Larrañaga v. Philippines, Communication No. 1421/2005, U.N. Doc. CCPR/C/87/D/1421/2005, ¶ 9 (July 24, 2006).

¹⁰⁰ See, *inter alia*, U.N. HUMAN RIGHTS COMM., Geniuval M. Cagas v. The Philippines, Communication No. 788/1997, U.N. Doc. CCPR/C/73/D/788/1997, ¶ 9 (Oct. 23, 2001); U.N. HUMAN RIGHTS COMM., Kurbonov v. Tajikistan, Communication No. 1208/2003, U.N. Doc. CCPR/C/86/D/1208/2003, ¶ 8 (Mar. 16, 2006).

information on the fate of a disappeared person and prosecution, trial and punishment of those deemed responsible;¹⁰² restitution of a victim's property;¹⁰³ grant of permission to leave the country;¹⁰⁴ issuance of a passport;¹⁰⁵ providing the victim with medical care;¹⁰⁶ and a guarantee that similar violations will not occur in the future.¹⁰⁷

The remedial practice of the HRC has been consolidated and is formally accepted by the majority of States parties. However, there are still some important challenges to be addressed by the HRC. Firstly, the HRC's practice on reparations has not been entirely coherent and systematic. Moreover, the ambiguity in the formulation of reparations and the underdevelopment of guarantees of non-repetition are problematic aspects that need to be rectified. Lastly, the HRC's development as described above has not been accompanied by the adoption of an adequate system to monitor the implementation of its requests for reparations.

Yet, against this background, two significant aspects should be singled out. First, it is notable how far the HRC's elaboration of reparation measures has developed in the absence of an explicit provision allowing it to consider reparations. This is also interesting in light of the difference between the terms "remedy" and "recours" in the ICCPR. In the English version, "remedy" is meant to embrace a procedural and a substantial meaning, while in French "recours" only has a procedural connotation.¹⁰⁸

¹⁰¹ See, *inter alia*, U.N. HUMAN RIGHTS COMM., *Rajapakse v. Sri Lanka*, Communication No. 1250/2004, U.N. Doc. CCPR/C/87/D/1250/2004, ¶ 11 (July 14, 2006).

¹⁰² See, *inter alia*, *Sarma v. Sri Lanka*, *supra* note 38, ¶ 11; U.N. HUMAN RIGHTS COMM., *Kimouche v. Algeria*, Communication No. 1328/2004, U.N. Doc. CCPR/C/90/D/1328/2004 ¶ 9 (July 10, 2007).

¹⁰³ See, *inter alia*, U.N. HUMAN RIGHTS COMM., *Simunek et al. v. The Czech Republic*, Communication No. 516/1992, U.N. Doc. CCPR/C/54/D/516/1992, ¶ 12.2 (July 19, 1995); U.N. HUMAN RIGHTS COMM., *Blaga v. Romania*, Communication No. 1158/2003, U.N. Doc. CCPR/C/86/D/1158/2003, ¶ 12 (Mar. 30, 2006); U.N. HUMAN RIGHTS COMM., *Victor Drda v. The Czech Republic*, Communication No. 1581/2007, U.N. Doc. CCPR/C/100/D/1581/2007, ¶ 9 (Oct. 27, 2010).

¹⁰⁴ See, *inter alia*, U.N. HUMAN RIGHTS COMM., *Celiberti de Casariego v. Uruguay*, Communication No. 56/1979, U.N. Doc. CCPR/C/13/D/56/1979, ¶ 12 (July 29, 1981); U.N. HUMAN RIGHTS COMM., *Saldias de López v. Uruguay*, Communication No. 52/1979, U.N. Doc. CCPR/C/13/D/52/1979, ¶ 14 (July 29, 1981).

¹⁰⁵ See, *inter alia*, U.N. HUMAN RIGHTS COMM., *Vidal Martins v. Uruguay*, Communication No. 57/1979, U.N. Doc. CCPR/C/15/D/57/1979, ¶ 10 (Mar. 23, 1982); U.N. HUMAN RIGHTS COMM., *El Ghar v. Libyan Arab Jamahiriya*, Communication No. 1107/2002, U.N. Doc. CCPR/C/82/D/1107/2002, ¶ 9 (Nov. 2, 2004).

¹⁰⁶ See, *inter alia*, U.N. HUMAN RIGHTS COMM., *R.S. v. Trinidad and Tobago*, Communication No. 684/1996, U.N. Doc. CCPR/C/74/D/684/1996, ¶ 9 (Apr. 2, 2002).

¹⁰⁷ This general formulation of guarantees of non-repetition can be seen in practically all communications where the HRC has found violations.

¹⁰⁸ ECKART KLEIN, *Individual Reparation Claims under the International Covenant on Civil and Political Rights: The Practice of the Human Rights Committee*, in STATE

One might question, for instance, whether the customary rule that prescribes reparation for the breach of an obligation would, in reality, be sufficient legal basis for the HRC to require reparation measures. However, the HRC has decided that it is article 2(3) (a) of the ICCPR that provides the legal basis upon which it can require States to remedy violations. Around 1983 the HRC rejected the idea that it was entitled to enforce its views, but stated that it could “nevertheless do something to bring [about] redress.”¹⁰⁹ According to the HRC, the preamble of the OP and article 2(3) of the ICCPR demonstrate that States parties intended the Covenant to be implemented, and therefore the HRC should indicate the remedies that a victim might benefit from.¹¹⁰

Second, the development of the HRC’s case law in this regard is particularly remarkable when compared to the practice of the European Court of Human Rights, which, despite the terms of article 41 of the ECHR, has decided to refrain from ordering any reparation measure other than compensation. The difference in the positions taken by the HRC and the European Court highlights that the option taken by the HRC can be commended and that the HRC has come closer to the IACtHR’s approach. Nonetheless, the IACtHR justifies its reparation requests on the basis of article 63(1) of the ACHR, while the HRC does so by relying upon the right to an effective remedy under the ICCPR.

B. THE ACTHR’S REPARATIONS AS A SOURCE OF INSPIRATION

The jurisprudence of the IACtHR on reparation is characterized by an innovative approach, based on an ample range of measures to provide a complete and detailed scheme of reparations, which often extend beyond the individual victim of a given case.¹¹¹ In fact, the IACtHR has identified different measures of restitution, compensation, rehabilitation, satisfaction, and guarantees of non-repetition. Among the diverse reparations granted by the IACtHR it is possible to find, for instance, requirements for the re-naming of street,¹¹² the building of monuments¹¹³ and other symbolic measures to the victims’ memory;¹¹⁴ the protection of human rights defend-

RESPONSIBILITY AND THE INDIVIDUAL. REPARATION IN INSTANCES OF GRAVE VIOLATION OF HUMAN RIGHTS 33, (Albrecht Randelzhofer and Christian Tomuchschat ed. 1999).

¹⁰⁹ DOMINIC MCGOLDRICK, *THE HUMAN RIGHTS COMMITTEE: ITS ROLE IN THE DEVELOPMENT OF THE INTERNATIONAL COVENANT ON CIVIL AND POLITICAL RIGHTS* 154 (1991).

¹¹⁰ *Id.*

¹¹¹ SERGIO GARCÍA, *LA JURISDICCIÓN INTER-AMERICANA DE DERECHOS HUMANOS: ESTUDIOS*, 163 (2006); CLAUDIO NASH, *LAS REPARACIONES ANTE LA CORTE INTER-AMERICANA DE DERECHOS HUMANOS (1988-2007)*, 37-78 (2d ed. 2009).

¹¹² *Myrna Mack-Chang v. Guatemala, Mertis, Reparations and Costs Judgment*, Inter-Am. Ct. H.R., (ser. C) No. 101 ¶ 286 (Nov 25, 2003).

¹¹³ *The 19 Tradersmen v. Colombia, Merits, Reparations and Costs Judgment*, Inter-Am. Ct. H.R., (ser. C) No. 109, ¶ 273 (July 5, 2004).

¹¹⁴ *Molina-Theissen v. Guatemala, Reparations and Costs Judgment*, Inter-Am. Ct. H.R., (ser. C) No. 108, ¶ 88 (July 3, 2004).

ers and personnel of the judicial system;¹¹⁵ the removal of waste from and reforestation of indigenous lands;¹¹⁶ the training of police, armed forces and other officials;¹¹⁷ and the re-opening of a criminal investigation.¹¹⁸

The jurisprudence of the HRC has progressively – though still modestly – evolved in a similar direction. It began by formulating general requests for States to undertake steps to provide victims with an effective remedy. It soon added a requirement to provide compensation, and subsequently it has also incorporated measures of restitution, rehabilitation, satisfaction, and a general suggestion of guarantees of non-repetition. Even though the HRC's reparations still lack the precision and breadth of those granted by the IACtHR, there has been important progress, which should continue.¹¹⁹

The previous sections discussed the significant influence exerted by the HRC's expansive interpretation of the right to an effective remedy on the IACtHR's jurisprudence regarding article 25 of the ACHR. Nevertheless, in terms of reparations, the influence might appear as flowing the other way. It is true that the HRC has not referred to the IACtHR's case law, but the potential influence of the regional court in this area cannot be underestimated. The broad understanding of the right to an effective remedy by the HRC has certainly had a considerable impact on the IACtHR's jurisprudence. But it is interesting to inquire as well whether, conversely, the IACtHR's decisions on reparations may influence the HRC to extend the scope of the right to an effective remedy even further. Whatever the impact that the IACtHR's jurisprudence on reparations may have had on the progress of the HRC in regards to the provision of redress, this influence should be increased. Looking at the practice of the IACtHR could help the HRC to articulate more comprehensive and detailed forms of reparations.

VI. CONCLUSIONS

The content and scope of the right to an effective remedy enshrined in article 2(3) of the ICCPR have experienced a significant expansion over time in the HRC's case law. This development seems understandable in light of the object and purpose of that provision, which was designed to provide an ample variety of mechanisms for the protection of the rights in the Covenant. Furthermore, article 2(3) had to serve too as a basis to request reparations that were not regulated elsewhere. Also, that expansive trend was built upon the initial decisions of the HRC in cases concerning

¹¹⁵ *Myrna Mack-Chang v. Guatemala*, Mertis, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 101 ¶¶ 270.b, 285 (Nov 25, 2003).

¹¹⁶ *The Kichwa Indigenous People of Sarayaku v. Ecuador*, Merits and Reparations Judgment, Inter-Am. Ct. H.R., (ser. C) No. 245 ¶¶ 293-95 (June 27, 2012).

¹¹⁷ *Fernández-Ortega et al. v. Mexico*, Preliminary Objection, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 215 ¶¶ 260, 262 (Aug 30, 2010).

¹¹⁸ *Nadege Dorzema et. al v. Dominican Republic*, Merits, Reparations and Costs Judgment, Inter-Am. Ct. H.R., (ser. C) No. 251 ¶¶ 248-9 (Oct 24, 2012).

¹¹⁹ See Valeska David, *Reparations at the Human Rights Committee: Legal Basis, Practice and Challenges*, 32 NQHR (2014) 8.

gross human rights violations, characterised by impunity. The interpretation of article 2(3) has evolved with such far-reaching consequences that this provision has practically become the source of all positive obligations of State parties. In this regard, it is worth noting that article 2(1) – which establishes the obligation to ensure the rights in the ICCPR – seems to have only a minimal impact in the resolution of individual communications. Almost every infringement of a positive obligation in relation to a substantive right is decided by declaring the violation of that right in conjunction with article 2(3) of the ICCPR.

The HRC has asserted that investigation, prosecution and punishment are, on the one hand, the effective procedural remedies to address grave violations of human rights. On the other, they constitute substantial remedies, that is, reparation measures. The HRC has outlined, however, that there is no right to demand that States prosecute and punish, with its support for a State duty to do so based on the broad individual right to an effective remedy. Nonetheless, it might be more consistent to justify an obligation without a corresponding right to prosecute and punish based on the aforementioned obligation to ensure the rights. Indeed, this idea according to which the duty to investigate, prosecute and punish derives from the obligation to ensure the right to life or to integrity was the previous position of the IACtHR, until it started referring to the HRC's case law and stretching the contours of the right to judicial protection in article 25.

It is in fact more difficult to understand the purposes and advantages of the shift made by the IACtHR in its interpretation of article 25 of the ACHR. The expansive interpretation of the right to an effective remedy by the IACtHR is related to the influence exerted by the HRC's jurisprudence. But in the Inter-American System the reasons and aims which gave rise to article 25 of the ACHR were different from those that resulted in article 2(3) of the ICCPR. Furthermore, unlike the ICCPR, the ACHR always had a separate provision for reparations. And the IACtHR, since the beginning of its work, developed a robust jurisprudence around the obligation to "ensure" the convention rights. Thus, the expansion of the right to an effective remedy has not only gone beyond the original ideas upon which article 25 was based and produced an amalgam of two different provisions of the ACHR (articles 25 and 8), but it has also given rise to further complexities, as seen above.

Notwithstanding how positive the recognition of a right to access to justice and to investigation, prosecution and punishment has been for the victims in the Inter-American system, the legal question of whether all that falls within the scope of the right to an effective remedy, would need to be re-examined. To improve the legal reasoning and argumentation of the IACtHR as well as to foresee counterproductive effects in this respect is also to reinforce the protection of human rights and the legitimacy of the Court.

Finally, an analysis of the development of the right to an effective remedy by the HRC cannot omit the fact that article 2(3) of the ICCPR has also been interpreted as providing the legal basis for the HRC's requests for reparations. Therefore, it is clear that this article fulfils multiple objec-

tives beyond those previously mentioned. By giving a substantive meaning to the right to an effective remedy, the HRC has extended the general request for compensation to include diverse forms of reparations such as restitution, rehabilitation, satisfaction and guarantees of non-repetition. Although the measures of redress provided by the HRC may not always be consistent and precise, the general evolution of the HRC's practice in this regard is remarkable. This practice shows that the path followed by the IACtHR can be a valuable source of inspiration for the HRC. Moreover, it is hoped that the way in which remedies are crafted and their domestic implementation could be improved in the near future. Perhaps in that way the broad reading of article 2(3) could eventually complete the development it has undergone with a view to fostering the realisation of human rights at the global level.

**BRITISH JOURNAL OF
AMERICAN LEGAL STUDIES**

VOLUME 3 - ISSUE 1

Spring 2014

Publisher: The British Journal of Legal Studies is published by Birmingham City University, Franchise Street, Perry Barr, Birmingham, B42 2SU United Kingdom.

Disclaimer: The Publisher, Editors, Members of the Editorial Board, associates and assistants cannot be held responsible for errors or any consequences arising from the use of information contained in this journal; the views and opinions expressed do not necessarily reflect those of the Publisher, Editors, Members of the Editorial Board, associates and assistants.

Copyright and Photocopying: The British Journal of American Legal Studies © 2013 Birmingham City University. All rights reserved.

Subscription Information: All UK enquiries should be addressed to the Law School, Birmingham City University, Franchise Street, Perry Barr, Birmingham, B42 2SU United Kingdom.

For all subscriptions in **North America**, please contact: Gaunt, Inc., 3011 Gulf Drive, Holmes Beach, Florida 34217-2199. Tel: 1 (941) 778 5211 or 1 (800) 942-8683; Fax: 1 (941) 778-5252. Email: info@gaunt.com

Information for Authors: We welcome the submission of unsolicited manuscripts for publication. Manuscripts should be electronically submitted via e-mail at BJALS@bcu.ac.uk. Citations should conform to *The Bluebook: A Uniform System of Citation* (19th ed. 2010). Further information can be found at <http://www.bcu.ac.uk/elss/school-of-law/research/centre-for-american-studies/british-journal-american-legal-studies>.

Citation: The British Journal of American Legal Studies should be cited as 3 Br. J. Am. Leg. Studies (2014).



British Journal of American Legal Studies
Birmingham City University School of Law
City North Campus, Birmingham, B42 2SU
United Kingdom

www.bcu.ac.uk/bjals