Prison Education: an interdisciplinary examination of prisons, punishment, schools, and education

Daniel Kalla

May 13, 2014

Contents

Introduction

An interdisciplinary examination of prisons, punishment, schools, and education, in the form of a Prison Studies and Restorative Justice Curriculum, and supplementary papers addressing prison statistics, punishment, synthetic thinking, and a language immersion education model.

I first came to Marlboro with an interest in education but no particular interest in prisons. In the spirit of trying things out I took a class (*Crime and Punishment*) about the prison system my first semester, and another class (*Race, Gender, and the Prison-Industrial Complex*) my third semester, so that by the time I had to decide on degree fields at the end of my sophomore year I was set on prisons and chose to create a "Prison Studies" degree field alongside Education. Since I didn't really have a strong background yet, I spent that summer reading through everything I could about prisons[1], while also teaching at an overnight Spanish immersion camp. Upon returning I continued to build my own Education and Prison Studies curriculum, formally in terms of classes and informally in terms of other community work. I served on the school's Community Court (official disciplinary body) for over a year and volunteered at the Brattleboro Community Justice Center, helping with a prisoner reintegration program (called Circles of Support and Accountability, or CoSAs). On the education end, I worked as the college math tutor for two years, taught at the Spanish immersion camp the following summer (I taught a variety of things in Spanish, including a month-long class about conflict resolution for high school credit), completed a TESOL certificate program (which included teaching English in Costa Rica as a practicum), co-taught a class on data visualization, and taught a class about the prison system and restorative justice my penultimate semester. I mention all of this to make it clear what my background is, and in particular the nature of my self-designed course of study.

Plan at Marlboro is intended to be an integrative, focused course of study, which is especially important for students majoring in multiple degree fields like me. Given my experience designing my own Prison Studies Curriculum and teaching a class about prisons, designing a Prison Studies and Restorative Justice curriculum seemed like the natural choice for a Plan project. Intended for a liberal arts college, this major is a twelve class course of study. The major structure includes four two-course sequences (conflict resolution and restorative justice; statistics and information literacy; oppression studies and poverty in the U.S.; and criminology and prison history), two electives, a class involving work in the community, and a thesis. In my paper, the seven goals of the major are detailed along with justification for why they're the goals, followed by the written curriculum. The curriculum portion includes a brief description of each course, learning objectives for the course, teaching methods (i.e. ways the teacher can help students accomplish each learning objective),

[1]Books I read include *Discipline and Punish*, John Conklin's *Criminology* (a textbook), Beccaria's *On Crimes and Punishments*, *The Culture of Fear*, *Madness and Civilization*, and *Life Without Parole: America's New Death Penalty?*

an annotated bibliography of course texts (mainly books, but also essential films and articles), and a breakdown of each course into different units, with descriptions. This curriculum mirrors my own course of study in certain ways (in particular, I made sure to read every resource that I decided to include), but departs from it in other ways (especially the formal class structure and intentional basis in mission and goals). As this part of my Plan is the largest and exhibits greater understanding of prisons and of education than any of my other components, this comprises 40% of my Plan grade.

The next two parts of my Plan further develop some of the major themes from the curriculum, with the goal of justifying it as a liberal arts curriculum. The first of these papers, *Prison Statistics, Invisibility, and Inequality*, examines the exclusion of prisoners from a variety of statistical measures and the effect of this, specifically in terms of concealing social problems and understating the level of inequality in the U.S. today. It aims to directly tie prisons to the understanding of other social problems, while highlighting the importance of statistical information literacy. Building on Becky Pettit's and Bruce Western's studies "invisible inequality" in the Current Population Survey (CPS) and the Prison Policy Initiative's studies of prisoners in the U.S. Census, my paper looks at social issues including non-employment, high school dropouts, voting, and population demographics. Throughout the paper I discuss the statistical methods used in a way intended to be understandable to a non-technical audience, as well as qualitative issues involved in these statistical analyses and sociological issues relevant to these statistical measurements. This paper comprises 15% of my Plan grade.

My third paper, *No Punishment Intended*, is a proposed program for restorative classroom management. This paper draws on principles of conflict resolution and restorative justice, while using academic theory surrounding inequality and the prison system to create a practical program. The literature review aspect of this paper examines the roots of zero tolerance policies in school and how they affect student development and learning, particularly focusing on theories of punishment in psychology and criminology. On the basis of this literature review it proposes a program with three parts, involving a method for dealing with disruptions in the classroom, basic conflict resolution pedagogy to teach students, and further conflict resolution pedagogy for teachers. I conclude by discussing some implementation issues and how this program would be assessed. This paper also comprises 15% of my Plan grade.

While the first three parts of my Plan were prepared with faculty feedback and guidance, the other three components were prepared independently. At the same time, they correspond to my different Plan advisers; since Education and Prison Studies are not normal majors at Marlboro I've done work for them through the Psychology, Statistics (Math), and Sociology Departments, and these three components loosely correspond to these three areas. As the purpose of the independent part(s) of Plan is to show a greater breadth of knowledge than is shown by the main part of the Plan, these parts do not relate as directly to education and prisons, though they do tie to other parts of my Plan. The first part, *Painting by Numbers*[2], is an examination in statistics and data visualization, meant to show more of my statistical abilities than I was able to show in Prison Statistics, Invisibility, and Inequality. I prepared a syllabus that described the structure of the exam and the topics I wanted covered, and then Matt Ollis designed the exam and gave it to me. This exam had two parts: a three-hour sit-down portion and a week-long take-home portion. During the sit-down part I was allowed use of the statistical package R, and not allowed to access

[2]The name for this paper comes from the data visualization class I co-taught with Matt as a second-semester junior.

the internet except through the R help system (and one of the exam questions that involved looking at online infographics), while I was allowed any non-faculty resources I wanted for the second part. In this Plan I include the syllabus I designed, exam questions and my responses from the first part of the exam, and then the questions and my responses form the second part. Aside from changing the font size, I have not modified any part of this exam since I took it in February. This paper is 10% of my Plan grade.

My second independent component, corresponding to sociology, is an ethnography of El Lago del Bosque – Bemidji, the Spanish immersion village I worked at the past two summers. This is the first time in my life I've written an ethnography (and one of the first times I've interviewed people for academic papers), so this is highly experimental. My main goal in writing this was to learn how to write an ethnography effectively, as well as to reflect on my experience working at the village (since I want to be a teacher, reflection is very important, and the majority of my teaching experience so far comes from the village). As with all ethnographies, it attempts to show the essential elements of life at the village while also using them to understand society more broadly: in my case, I use El Lago del Bosque as a lens for examining language and gender. This is based in my personal experience as a counselor and villager, and supplementary interviews of staff members from this past summer (2013). I ultimately decided not to publish this with the rest of my Plan, so it is omitted here. This paper is 10% of my Plan grade.

My final paper, *Synthetic Thinking – A Manifesto*, promotes a particular type of thinking that served as a sort of guiding philosophy in my Plan, and corresponds loosely to Tom Toleno in psychology. While the curriculum at the beginning of my Plan lays the basis for most of my other papers, this essay ties together major aspects of each of my other components into a single argument. It therefore works as a sort of conclusion to my Plan. After defining synthetic and critical thinking, this paper details some of the shortcomings of a focus on critical thinking, promotes the value of synthetic thinking, and provides a model for using the two types together in order to decide the truth of a given claim. While I use the examples of statistics and feminist critical theory throughout the paper, I also draw examples from fields as diverse as business, conflict resolution, political theory, and systems thinking, among others. This paper is 10% of my Plan grade.

As a final note before the main content of my Plan, I want to thank the communities of Marlboro College, Marlboro College Graduate School, and El Lago del Bosque Bemidji for all the help and amazing experiences they've given me. Special thanks to Tom Toleno, Matt Ollis, and Kat Rickenbacker, my Plan sponsors, as well as John Sheehy, Renee Byrd, Ken Schneck, and Jerry Levy, who are not official Plan sponsors but all contributed greatly to my Plan. Most of all, of course, I want to thank my parents, Anne Krisnik and Mark and Christine Kalla, for bringing me up and supporting me throughout college and life. You're all great, and you deserve a better dedication than I have time to write before my Plan is due. Thank you so much, I couldn't have done this without you.

Part I

Prison Studies and Restorative Justice Curriculum

Chapter 1

Mission and Goals for a Prison Studies and Restorative Justice Curriculum

Executive Summary

This document is an introduction to the Prison Studies and Restorative Justice curriculum; it explains the overall goals of the curriculum and points to what parts help meet each of those goals. This document is inteded for internal department use, to help them assess how well the curriculum fulfills the mission and meets the goals and to guide them through curriculum revisions.

Mission Statement

This curriculum prepares students for community-oriented, theory-informed work in fields involving prisons, restorative justice, inequality, and conflict.

Introduction

This curriculum is for a "Prison Studies and Restorative Justice" major at a four-year college or university, with this document detailing the overall goals of the curriculum. Goals are different from learning objectives in three main ways. First, goals provide a sort of guiding vision in designing (and re-designing) the curriculum and are often more useful when somewhat broad of vague, while learning objectives are used in part to measure the success of the curriculum and must be specific and measurable. Second, learning objectives are strictly concerned with academics, while goals can be much broader (goals 5 and 6 are particularly good illustrations of this). Third, learning objectives are normally used are requirements for students passing the course, while goals are not. As learning objectives are used at the level of individual courses (and students), goals inform the whole curriculum. Finally, goals can differ from learning objectives because they don't need to be measurable to the same extent. While it is important to find ways to measure goals, for this structure goal 6 ("Students will be involved in improving their communities") is more useful than

a more measurable goal like 'students will spend at least one year doing community learning,' because with goal 6 it is worth working to increase community involvement whether students are doing community learning or not. There'smore to community work than can be captured in a single measurement, so while measurements are important to track progress, at a higher level goals may be more useful when they're broader.

The main reasons goals are so essential even in a static curriculum is that they inform institutional structure and what things the department should put resources toward supporting. To use goal 5 ("Students will be lifelong readers") as an example, a department should put substantial resources toward building a strong departmental library and advocating for the school librar in general to encourage reading. Faculty should actively promote reading by modelling the behavior as avid readers themselves, by organizing book clubs and recommending books to students on a regular basis, etc. (They are by no means required to do these specific things, but they are expected to work to promote the curriculum goals beyond merely teaching courses in a way consistent with the syllabi.) Having a good understanding of goals is even more important when the curriculum is being reformed, because some parts of it may not make sense in terms of the rest of the curriculum. Probably the best example of this in the mandatory First Aid/CPR training, which makes minimal sense from a curricular standpoint but contributes to goal 6 ("Students will be involved in improving their communities") by making sure they have the skills to act (and potentially save lives) in an emergency.

With these distinctions in mind, the seven goals of this curriculum are below. ("SWBAT" is used as a shorthand for "Students will be able to", and similarly for "SWB".) There is a discussion about the rationale for each goal chosen, along with a descriptions of what parts of the curriculum most directly address that goal.

1.1 SWBAT understand the history and state of prisons and restorative justice

This goal covers the essential content knowledge of the curriculum that makes it a major.

1.1.1 SWBAT understand the history and state of prisons

Prisons are useful to study because they have strong connections with nearly every major social issue (e.g., racism or anything dealing with inequality, poverty, unemployment, crime, and violence). Foucault talks about how prisons encapsulate a number of important changes in western society [57, p. 115; 121-5] and how they are tied to most other major social institutions [60, p. 301-2]. In addition, Becky Pettit, Bruce Western, and others talk about how prisons distort the dominant statistics used for assessing and making policy, so that an understanding of prisons is essential to understanding many social issues [36, 126, 130, 176].

Support in Curriculum The two-course prisons sequence ('Introduction to Criminology' and 'History of Prisons') is the largest part of the curriculum focused on this goal, although most/many of the electives are prison-focused also. In addition, the poverty course sequence ('Introduction to Oppression and Identity Studies' and 'Poverty in the U.S.') addresses inequality and systems of oppression, which is an essential part of the understanding the role of prisons today. Goal 1.1 and/or goal 1.2 will also be addressed through the senior thesis.

1.1.2 SWBAT understand the history and state of restorative justice

Restorative justice is an important part of understanding the current prison landscape, while also being useful as an object of study because it addresses isssues of conflict resolution more broadly. In particular, it focuses on group relationships, inclusion, and healing, which retributive justice (the dominant model in the U.S.) sorely neglects [184, p. 22-4]. In general, it effectively addresses the weakness of retributive justice and is important in terms of thinking of alternatives in the justice system [184, p. 59]. While it started in the context of criminal justice, it has a lot of usefulness in schools, business, and other places because it deals with issues essential to any community [129, p. *xiii*].

Support in Curriculum The two-course Restorative Justice sequence ('Introduction to Conflict Resolution' and 'Restorative Justice Practices in the U.S.') is devoted specifically to restorative justice. The poverty course sequence is also relevant here, and the 'Introduction to Criminology' class provides some broader theoretical background (since the restorative justice focused is relatively skill-based and empirical). Community work may incorporate restorative justice, as well. In additon, the "Community Work and Reflection" class and the "Senior Thesis" class will make heavy use of different circle processes that form a key part of restorative justice (and promote communication and community-building).

1.2 SWBAT resolve interpersonal conflicts in their lives

Interpersonal conflict is inevitable in all aspects of life, and so conflict resolution is an essential skill for work, family life, and getting involved in the community. Stephen Covey talks about how the skills involved in resolving interpersonal conflicts are powerful in resolving all types of problems [39, p. 7-8], but they also prove worthwhile on the basis of interpersonal conflicts alone. As an example, in *Getting to Yes*, Fisher and Ury talk about the disadvantages of "positional negotiation", which is the dominant style in the U.S. In particular, it leads to worse outcomes, is less efficient, endangers relationships, and is essentially impossible when there are more than two parties [59, p. 7-8]. "Soft" negotiation, which most people see as the alternative, results in exploitation by the other party [59, p. 10]. For that reason they promote the alternative of "principled negotiation" [58, p. 13], although there are a number of other methods of conflict resolution available (such as nonviolent communication) [137]. This curriculum will therefore try to teach multiple methods of resolving conflicts, so that students find one that they are able to use effectively. In addition to being an essential part of practicing restorative justice (and a skill many offenders need to learn, and may need help mastering), conflict resolution is a general life skill that everyone should have.

Support in Curriculum This is addressed most directly in the 'Introduction to Conflict Resolution' course, although the skills are built upon in 'Restorative Justice Practices in the U.S.'. In addition, clear, nonaggressive verbal communication is emphasized throughout the curriculum, and conflicts that arise during community work are discussed in the linked class.

1.3 SWBAT use academic theory to address social problems in real life

In *Pedagogy of the Oppressed*, Paulo Freire talks about how praxis (theory-informed practice) is essential for creating social change [62, p. 51]. Being able to connect theory to real life makes it useful, makes social change possible, and promotes learning because it makes the theory more intrinsically interesting. This is an essential skill to focus on because of anti-intellectualism outside of schools [71, p. 97-101] [6], and even within them [71, p. 94-7]; if a school doesn't focus on applying theory many students may never learn how. In addition to helping students promote social change, this is important (for social science in particular) because people cannot fully understand a theory without seeing how it connects to reality. Their understanding will be very narrow or wrong (sometimes both). Being able to connect theory to real life improves the experience of both.

Support in Curriculum The most important part of the curriculum in terms of addressing this goal in the community service-linked class, in which students are explicitly required to draw parallels between theory and their actual experiences. Most students will integrate their community experiences into their thesis, although this isn't an explicit requirement. Before the community service portion of the curriculum, students should be able to use academic theory from the conflict resolution class in their lives, and both that class and the Introduction to Identity and Oppression Studies class will involve students reflecting on their personal experiences. Other classes may not as directly promote this goal, but teachers can still work to encourage it.

1.4 SWBAT research and use history to understand problems more deeply

Connected with being a lifelong reader, students need to be able to find new information and to be able to use that information to make theory. Information literacy is essential for ensuring their continued development as scholars, and important because no academic program can possibly teach them everything it would be useful for them to know, in particular because many of them will later have jobs that don't exist yet [181]. According to the Association of College & Research Libraries, "Information literacy forms the basis for lifelong learning. It is common to all disciplines, to all learning environments, and to all levels of education. It enables learners to master content and extend their investigations, become more self-directed, and assume greater control over their own learning" [9]. They note a number of skills it includes, in particular the ability to "[i]ncorporate selected information into one's knowledge base" and "[u]se information effectively to accomplish a specific purpose." [9] So students should be able to build and use a knowledge base for the future. Together with the second goal, this effectively implies that they will be lifelong learners.

Support in Curriculum The information literacy sequence ('Introduction to Statistics' and 'Research and Information Literacy') directly tries to build students research and information literacy skills. There is a strong emphasis in the rest of the curriculum on history and context, and the senior thesis requires a significant research component also.

1.5 SWB lifelong readers, not just during college

Reading is one of the single habits most predictive of success, as well as personal happiness. Avid, skillful readers have myriad opportunities available to them, while people who struggle to read have severely restricted life opportunities. In addition, reading constantly exposes people to new ideas, keeps them always learning, and is consistently considered personally fulfilling by readers. It correlates closely to declarative knowledge after controlling for measures of general ability, and it predicts verbal skills even more strongly [43, p. 144; 143]. Even for children, many experts believe that reading volume is more important than oral language in causing differences in verbal skills [43, p. 138], and this would almost certainly be truer for adults. However, it's worth noting the importance of verbal skills from oral language: amount of vocabulary children are exposed to is one of the most important factors in explaining all types of academic success later in life [147]. Since many (probably most) of the graduates of this program will have children later in life, their own verbal skills are important in affecting their children's verbal skills.

Support in Curriculum This goal is probably the hardest to support by curriculum structure, but not less important as a result. Students are required throughout their classes to read quite a bit, and for the electives they must choose books of interest for themselves. Other than the electives, this curriculum will encourage reading by having teachers model the behavior, exposing students to a lot of good books (that they are required to read), and advocating directly for a strong school library.

1.6 SWB involved in improving their communities

People who have the privilege of a college education have an opportunity to affect the world that other people do not, and a corresponding obligation to help. Failing to work to improve their communities means that a college education functions to promote social stratification (Strong 51). By improving their communities, graduates give back to societies whose structures have given them educational opportunity in the first place, avoid entrenching inequality, and work toward positive, necessary social change. Active community involvement also fulfills the original function of schools to create an informed, politically active citizenry necessary for the basic functioning of democracy.

Support in Curriculum The biggest contributor to this goal is the mandatory year of community work, although other community service opportunities will also be promoted. (There is also the the First Aid/CPR requirement (because certification is easy and potentially lifesaving), which gives students skills to potentially contribute to any type of community.) This goal also ties into goal two and the poverty/inequality sequence (and most of the rest of the curriculum in some way); by giving students the skills to resolve conflicts effectively they will automatically contribute more to their communities, and by teaching them about issues related to poverty etc. they will become more interested in addressing social problems. Obviously not all students will become substantially more interested and involved in addressing major social problems, but large number of them probably will.

1.7 SWBAT describe and make use of their academic strengths

As college is a place for people to figure out who they are, they should also learn what their academic strengths are and make use of them. This makes them the most effective they can be in academics, community work, and other parts of their lives. Rath talks about how, regardless of whether anyone can learn to be anything, people are much happier and more effective when they use their strengths [132, p. *i-iv*]. In addition, most college students are in late adolescence and transitioning to adulthood. Discovering one's identity and individuation are huge parts of this transition and include discovering and developing one's strengths [100, p. 94, 103].

Support in Curriculum This goal is somewhat hard to tie to specific parts of the curriculum, because the largest part of this will be teachers focusing on students strengths in terms of feedback, activities, etc. That said, this curriculum tries to keep a significant balance between skills and theory (or praxis and theory), between qualitative and quantitative analysis/data, and other areas, in addition to crossing many disciplines and exposing students to their methods and philosophies. The elective structure (while a bit limited) is also intended to help students develop their particular areas of relevant strengths (and interests), and of course the community work and senior thesis are highly personalized and based on students' individual strengths.

Chapter 2

A Curriculum for Prison Studies and Restorative Justice as a College Major

Abstract

This curriculum describes all the courses, course objectives, teaching methods, prerequisites, book list, and units for each class in the curriculum (with the electives being less detailed). The structure of the curriculum is as follows: four two-course sequences (in conflict resolution and restorative justice; information literacy; criminology and prison history; and inequality and poverty), followed by two electives, a course supporting some type of community service, and a senior thesis.

Introduction

This curriculum reflects the rise in importance of prisons and restorative justice in the U.S., and the lack of college majors that address them. Prison Studies is a major unique to Marlboro College (and lacking in curricular structure), with the next closest major elsewhere being criminology or prison mangement, or perhaps sociology with a specialization on prisons. Restorative Justice exist as a major elsewhere, but still not in a way that's tied directly to studying prisons broadly the next two closest majors, Fresno Pacific University's "Criminology and Restorative Justice Studies" and Lipscomb's "Restorative Criminal Justice", are much more focused and are closer to technical majors. Prison Studies and Restoartive Justice, as a college major described here, is a part of the liberal arts tradition. Rather than specialize and focus on training for specific careers, this major gives people an understanding of issues like resolving conflicts, inequality, and information literacy, as well as some specific content in prison history and restorative justice practice. To this end, there are four two-course sequences included in this curriculum, focused on restorative justice, information literacy, prisons/crime, and inequality. In place of a practicum preparing someone for a career, there is a community service project that prepares students for future community work of interest. Also in the liberal arts tradition, electives for this major are more wide-ranging than those that exist in similar majors, so students can study broadly and focus on what they find the most interesting and relevant. Because this major combines prisons and restorative justice as objects of study (a fairly natural pairing), and does so in the context of a liberal arts education, this program

is unique and able to contribute to students' options.

As the first student to complete a major in Prison Studies, a volunteer in Vermont's restorative justice system, and someone experienced teaching, I am uniquely situated to design this curriculum. I've taught a college-level class about prisons and restorative justice and a high school class about conflict resolution, as well as co-teaching a class about data visualization (a type of statistical information literacy). This is not to say that I can design it *better* than anyone else would be able to, but my position is unique. Moreover, since other curricula like this do not yet exist, my interest and willingness in creating it mean I m the person best situated to do so. Given that professors generally design their own classes this is intended more as a guide than anything else; while the broad subject matter and learning objectives of the courses are important to keep for the shape of the curriculum as a whole, individual teaching methods and course texts can change based on individual teachers' interests and strengths. Teaching methods, units, and course texts are listed only as guidance, to assure colleges that this curriculum can be implemented and to help individual teachers with ideas. A teacher struggling to accomplish a given learning objective has the listed teaching method available as one possible tool, to use if they find it helpful and to ignore otherwise. The situation is similar for course texts: while an experienced teacher may texts they prefer or are best able to teach with, they have these as options to fall back on. I have read through all of the listed texts in full and know that they are solid choices[1], even as better options may be available. More than anything, these mini-syllabi reflect how *I* would teach these courses, which is the best starting point I can design for implementing this as a college major. The courses of this curriculum are below.

[1]There is one excpetion: I haven't read all of *Women's Voices, Feminist Visions* (recommended reading for one of the classes) because it's a anthology. I have read about half of it, which is enough to justify it in general, and even more so for recommended reading.

2.1 Introduction to Conflict Resolution[2]

After developing some basic skills in communication, this course examines different theories of conflict resolution, the nature of conflict, and the mediation process. It uses real and simulated conflicts to practice skills as well as to understand conflict more deeply. Although this is intended to be a general interest course, it should provide a foundation for restorative justice work in terms of theory and practice of conflict transformation.

Course Objectives

Students will be able to

1. demonstrate communication skills in reflective listening, assertion, and conflict resolution.
2. describe and apply the theory of principled negotiation to simulated conflicts.
3. describe the causes of conflict, as well as its positive and negative effects.
4. mediate a simulated conflict, and decide when conflicts should be resolved through mediation versus negotiation.

Teaching Methods

Teacher will give students the tools to accomplish these goals by

1. breaking down reflective listening, assertion, and conflict resolution into concrete, formulaic skills.
    - Reflective listening: students listen to other students/teacher and try to explain exactly what they said (i.e., the content of what they said, and the emotion behind it). In addition, for this and assertion, students practice expressing and categorizing emotions so they can be more specific when they reflect feelings.
    - Assertion: teachers explain how to make three-part assertion messages using examples, then have the students practice translating other language into three-part assertions. Students use the charts of emotions from reflective learning.
    - Conflict resolution: teacher explains Bolton's model for conflict resolution, then students practice applying it to simulated conflicts. (This also provides a review of reflective listening and assertion skills.)
2. describing the five major parts of principled negotiation[3], then having students practice negotiating in simulated conflicts.

[2]The term "conflict transformation" is a common alternative to "conflict resolution" because conflict resolution sees conflict as bad and avoidable while conflict transformation views it as unavoidable and sometimes positive or necessary. While philosophically I agree with that and include *The Functions of Social Conflict* to address the positive aspects of conflict, ultimately this class is primarily focused on resolving conflicts. It serves as a prerequisite for the class about restorative justice, which is essentially a type of conflict transformation.

[3]Avoiding positional negotiation; seeing the other's needs; looking for creative, mutually beneficial solutions; using objective criteria; and considering one's BATNA (Best Alternative to a Negotiated Agreement)

3. explaining the functions of social conflict, and having students generate examples of conflicts illustrating different functions, including what causes these conflicts and what negative effects they have. (Students may be encouraged to use earlier example conflicts from class or conflicts from their own lives.)

4. explaining the structure of a successful mediation and having students practice this with simulated conflict. For a conflict in class, the students in the simulated conflict can practice good or bad communication or negotiation, while mediators can practice mediating (and they people switch roles for the next conflict).

Outline of Course Topics

Communication and Conflict Resolution This unit introduces the concept and purpose of conflict resolution and works to develop basic skills in communication (in particular: active listening, non-verbal communication, assertiveness, and a model for conflict resolution). Class mainly involves discussing the readings and doing communication exercises; this unit is heavily focused on skill development. Homework involves readings from Bolton, as well as reflections on exercises in class.

Principled Negotiation This unit focuses on the text *Getting to Yes* and the theory of principled negotiation. Class will mainly involve discussion of the readings and simulated conflict in which people practice the different parts of Fisher and Ury's framework. Other than reading the book, homework mainly involves exercises (e.g., find a creative solution, determine your BATNA, etc.), as well as short responses and a short essay responding to questions in the book.

Purpose of Conflict and Conflict Transformation This unit looks at causes and functions of conflict, focusing on when conflict resolution is worthwhile and how to do it more effectively. In particular, this unit examines some of the benefits of conflict and why it might be necessary, so that students can better resolve conlicts in a way that preserves as many of their positive qualities as possible (or decide that they do not need to be resolved). The unit is based primarily on the text *The Functions of Social Conflict*, and classes mainly involve discussing the text and applying it to conflicts discussed in previous units. Homework mainly involves reading and reflection pieces.

Mediation This unit brings together everything that has been learned in previous units and focuses on developing the skills of mediation. Students read through *The Mediation Handbook* and practice mediation skills in the context of simulated conflicts. Class is focused on mediating simulated conflicts, along with supplementary discussions and clarifications of material, while homework mainly involves reading the handbook and writing reflections about the simulated conflicts in class. The class will terminate with mediation of a major simulated conflict (or more than one, so that students can choose), and for the final paper students will write about this conflict and the mediation process.

Course Text(s)

***People Skills: How to Assert Yourself, Listen to Others, and Resolve Conflicts* by Robert Bolton [22]**

This book focuses on the communication skills necessary for successful conflict resolution, making it an excellent introductory text. It starts by talking about communication generally, moves to listening skills, discusses assertion skills, and ends by considering conflict resolution. The book is clear, thorough, and reasonably concise, and focuses on skills in terms of specific behaviors.

***Getting to Yes: Negotiating Agreement Without Giving In* by Roger Fisher and William Ury [59]**

Fisher and Ury's *Getting to Yes* is a classic book on conflict resolution, and their theory of "principled negotiation" is one of the most influential theories in conflict resolution today. The book is short, clear, and very readable, has numerous examples of successful conflict resolution, and has a business focus that complements the course's (and curriculum's) general focus on crime. Its organization into chapters focusing on discrete skills and practices means it fits well in a classroom and that class sessions can help students practice and master the principles described in the book.

***The Functions of Social Conflict* by Lewis A. Coser [38]**

This 1956 book details a number of sociological benefits to conflict, in particular looking at how it can increase social cohesion. The book is mainly based in sociological theory (in particular, the work of George Simmel), but it also includes examples of various functions of conflict. It helps students decide when conflict resolution is necessary or helpful, while also giving them deeper insight into the causes of conflict and helping them resolve conflicts in a way that does not interfere with their positive functions.

***The Mediator's Handbook* by Jennifer E. Beer, Caroline C. Packard, and Eileen Stief [17]**

The Mediator's Handook provides an in-depth look at how to do mediation, both in terms of logistics of setting up a session and in terms of developing the skills of an effective mediator. The authors are all experts in the field and the book is full of illustrative anecdotes from their experiences. It provides the structure in a way useful for beginners and offers more detailed, nuanced advice for experienced mediators. For that reason, it is effective both as a starting book and as a reference book, meaning its use in this class fits in with the overall curriculum of the major quite well.

Recommended Reading

***From Dictatorship to Democracy: A Conceptual Framework for Liberation* by Gene Sharp [144]**

This book details the process of effecting nonviolent revolution, specifically for overthrowing dictatorships. As an introduction to conflict resolution, this class is interested in looking at not only how conflicts can be mediated or negotiated but also how conflict can be used to effect necessary social change. The book provides a framework for enacting nonviolent revolution and talks about

when negotiation or other normal conflict resolution strategies are *not* appropriate. This book goes with Lewis Coser's *The Functions of Social Conflict*; they work together to help students define when conflict resolution is appropriate (as opposed to "conflict transformation").

The text is available free online, but there's no direct link (search 'From Dictatorship to Democracy pdf' in Google).

2.2 Statistics for Social Sciences[4]

This course in an introduction to statistics, covering issues in understanding data, inference, linear regression, and analysis of variance. Important qualitative issues are integrated into the teaching of basic statistical concepts, and students end the course by focusing important aspects of evaluating others' use of statistics (in media and academia). Somewhat unusually for an introductory course in statistics, this class uses nonparametric statistics as the basis for statistical inference, both because it is pedagogically easier and because it makes more sense philosophically.

Course Objectives

Students will be able to

1. understand essential concepts in statistics like sampling, significance, variance, etc.
2. test hypotheses using statistical inference.
3. perform and interpret linear regressions.
4. know when the use of statistics is appropriate to answer a research question.
5. critique others' improper uses of statistics.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. explaining the concepts and having students explore datasets in a way that uses these concepts.
2. showing them the process and give them problems to solve.
3. showing them the process and give them problems to solve and exercises in interpreting regressions.
4. giving students research questions and having them decide what type of data they need.
5. giving students examples of statistics and having them critique others' (related to 1 and 4 above).

[4]I use this name despite my criticism of the idea of "Statistics for Social Sciences", as though the material in those classes should not be the material in almost all introductory courses. Natural scientists face most of the same qualitative issues that social scientists face, and in natural sciences these are more often ignored. Nevertheless, I use this name because qualitative issues are especially important in the social sciences and I don not want another department to start a competing "Statistics for Social Sciences" course.

Outline of Course Topics

Types of Data and the Normal Distribution This unit will look at different types of data and characteristics of the normal distribution. As the first unit it will also introduce most of the vocabulary and basic concepts of the term. *Introduction to Statistics* Chapters 1-5 ('Introduction', 'Organizaing and Summarizing Data', 'Intuitive Inference', 'Probability', and 'The Normal Distribution'); *How to Lie with Statistics* Chapters 1-3 (Sample Bias, Measures of Center, and Small Samples and Badly Specified Data). Homework mainly includes reading and selected exercises from the textbook. Students should also learn the basics of the programming language R in this unit.

Statistical Inference The unit will examine the foundation of statistical inference and the hypothesis testing framework, after which point it will focus on developing the skills of statistial inference in t-tests and ANOVA tests. *Introduction to Statistics* Chapters 6-12 ('Hypothesis Testing', 'The Wilcoxon Two-Sample Test', 'Nonparametric and Parametric Tests', 'Estimation The Two-Sample Shift Model', 'Point Estimates, Confidence Intervals, and Tests of Hypotheses', 'The One-Sample Problem', and 'The Two-Sample Problem: Paired Observations'); *How to Lie with Statistics* chapters 4 and 8 (Hypothesis Testing and Correlation vs. Causation). Homework mainly includes reading and doing statistical inference (mix of textbook exercises of real-life examples); students should be gaining a basic proficiency in R.

Linear Regression and Analysis of Variance The unit examines linear regression of data, including foundations of regression, hypothesis testing, and modifications of data (e.g., taking the logarithm), as well as analysis of variance. *Introduction to Statistics* Chapters 13-18 ('The ANalysis of Bivariate Data', 'Least Squares Regression and Correlation', 'Comparative Experiments: k-Samples', 'Analysis of Variance', 'The Analysis of Categorical Data', and 'Chi-Squared Tests and Two-Way Classifications'), *How to Lie with Statistics* chapters 5-7 (Displaying Data on Graphs (both 5 and 6) and Proxy Variables). Homework mainly involves reading and selected exercises.

How and When to Use Statistics This unit will look at the process of gathering information with statistics and how to interpret other people's published statistics. It will involve many examples of published uses of statistics as well as various research questions, for which students will decide what data would appropriately answer that question. *How to Lie with Statistics* chapters 9-10 ("How to Statisticulate" and "How to Talk Back to a Statistic"), *Why Most Published Research Findings are False*, various examples of uses of statistics. Homework includes a bit of reading, but other than that mainly includes research exercises and writing critiques of statistical presentations and claims. The class will end with a comprehensive take-home exam.

Course Text(s)

***Introduction to Statistics: The Nonparametric Way* by Gottfried E. Noether [121]**

This book starts with an examination of statistical data and then introduces statistical inference from the perspective of nonparametric statistics. It ends by discussing linear regression and analysis of variance. This book is clear and well-written; the computer commands are outdated since the book is old but otherwise it's quite strong. This book is chosen primarily for its strong discussion of issues in data analysis.

***How to Lie with Statistics* by Darrell Huff [88]**

This book gives a broad overview of common ways people misuse statistics, with an easy, readable style and many real-life (if old) examples. In addition to helping people learn how to interpret and use statistics, this perspective can help clarify the meaning and importance of central statistical concepts. While it's a short read, here I use it as a textbook so that it matches up with the units and so that students regularly have interesting readings outside of the textbook.

***Why Most Published Research Findings are False* by John P. A. Ioannidis (article) [91]**

This article makes use of the students' statistical understanding to make a strong case for being critical of published statistics, while also detaling many of different factors that cause statistical conclusions to be more or less reliable. A side benefit of this article is that it focuses more on natural sciences than social sciences to make its case, which makes it more effective overall given the greater trust placed in natural sciences.

2.3 Introduction to Identity and Oppression Studies[5]

This class will also almost certainly exist in another department, either sociology or Gender Studies. If not, though, the course I design would be based more in Gender Studies than anywhere else because the scholarship is well-developed, even as the course is more focused on looking at oppression generally.

Course Objectives

Students will be able to

1. explain what a social construct is and the sociological importance of race, gender, class, and disability.

2. explain (with examples) how race et al. are embedded in our social structures, and the connections between privilege, dominance, and oppression.

3. detail the history of disability in the U.S. and how it relates to other axes of oppression, using the concept of intersectionality.

4. discuss the importance of their own identities and the communities they are a part of in their lives.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. defining social construct; give examples of how these meanings have changed throughout history; students share personal experiences.

2. giving some examples, and have students generate their own (again using lived experience when possible).

3. presenting the history of disability in the U.S. and asking students to explains how it's tied to other forms of oppression.

4. modeling with their own autoethnography and discussing different communities in the context of Anzaldúa

[5] "Identity and Oppression Studies" is used as a stand-in title for a number of possible phrases, most likely "Gender Studies", "Race Studies", "Women's Studies", etc. The choice of the term "Identity and Oppression Studies" also reflects my belief that separating race, gender, class, etc. studies at an introductory level makes little sense, and has ever since intersectionality become a dominant paradigm in the field. It is uncommon to try to integrate different identities and the idea of intersectionality into an introductory course, but I contend that it is because it is only talked about later (and not from the beginning) that people find it difficult and struggle to think in fully intersectional terms at higher levels. "Identity Studies" might plausibly be a better name for the course in general, although given the focus on power and privilege, oppression is an essential part. In a Prison Studies and Restorative Justice Curriculum, race, gender, etc., are important almost fully because of their connection to oppression, so "Oppression Studies" is particularly approrpiate.

Outline of Course Topics

Introduction to Identity and Oppression Studies This unit will introduce students to the basic ideas in feminist theory, including the ideas of a social construct and of intersectionality. The goal is to familiarize students with the idea of looking at society from a feminist perspective and help them understand why feminism can be useful and relvant to their own lives. *Feminism is for Everyone: Passionate Politics* and *Crash*, as well and supplementary articles foundational in feminist theory. Assignments mainly involve reading and reflection pieces about students' own lives. (In this unit, the instructor also discusses the choice of "Identity and Oppression Studies" as the title of the course over "Gender Studies" or something similar.)

Power, Privilege, Dominance, and Invisibility This unit looks at how power is exercised over different groups, the invisibility of privilege, and how this is related to dominance and oppression. Specifically, it looks at this in terms of surveillance (which is becoming increasingly important to understand our society) and by using the example of a man who physically changed his race in 1959 and exprienced the difference in privilege firsthand (this talks about race, but class discussion will work to connect it back to oppression more generally and incorporate intersectionality). *Black Like Me* (book and film) and *SuperVision: An Introduction to the Surveillance Society.* Homework for this unit involves readings and reflective writings looking for instances of power and privilege in their own lives.

Disability: Another Axis of Oppression This unit looks at disability, which is often ignored in identity and oppression studies[6], and uses it to better understand the history of oppression in the U.S. in terms of race, gender, class, and other dimensions. Because it discusses how different forms of oppression are related, this unit asks students to consider the importance of intersectionality more explicitly than previous units. *A Disability History of the United States*, along with supplementary articles talking about oppression in history. Homework mainly involves reading and reflection pieces.

Borders Between Cultures This unit examines borders of all types between cultures, and the identity of people who live on these "Borderlands". It also asks students to formally reflect on their own identity and communities, and how these are important in their lives. *Borderlands/La Frontera: The New Mestiza.* Homework involves reading and working on the final project, an autoethnography. This project requires students to examine their own lives and identities, how they formed, and how they affect their lives. By default the project is a paper, but students are stongly encouraged to do whatever helpes them best express their identities. As *Borderlands/La Frontera* uses poetry and a mix of language, this unit will also express different ways of knowing and communicating knowledge academically, and it is hoped that Anzaldúa will help inspire students in their final projects. At the end of class, students will present their projects to the class before turning them in.

[6]In general, identity and oppression studies confines itself to sex, race, and class, as well as issues directly related (e.g., gender, ethnicity, language)

Course Text(s)

***Feminsim is for Everybody: Passionate Politics* by bell hooks [83]**

This book provides an accessible entry to feminsim and is written by a major theorist in the field. It discusses major events in the history of U.S. feminism and looks at how the movement connects with different aspects of daily life. Supplemented by major articles in feminist theory, this book is a strong introduction to the central ideas of identity and oppression studies.

The book is also free online:

`http://excoradfeminisms.files.wordpress.com/2010/03/bell_hooks-feminism_is_for_everybody.pdf`

***Crash* by Paul Haggis [79]**

This film looks at the interconnected lives of a variety of characters from different races, classes, and genders. The idea of intersectionality is central to the film, as well as the idea of complex personhood (the idea that people have both positive and negative traits, and act differently in different situations (in other words, generalizations about people fail to hold even at the individual level)).

***Black Like Me* by John Howard Griffin [75]**

Black Like Me is a classic text describing the author's experiences changing races (becoming visually black) in the South during the ending years of Jim Crow. Because he had a unique opportunity to experience anti-black oppression, he gained a lot of insight into how he was privileged as a white man, and seeing how race was relevant in areas he had never thought about previously. This book will be a useful text to discuss in terms of defining privilege, looking at the centrality of race to various social structures, and thinking about race and other aspects of identity as fundamental or immutable qualities. It also complements the rest of the course by acting as a fairly specific example of a type of oppression (specifically, racial oppression), but is short enough that intersectionality still definitely dominates the course.

***Black Like Me* by Carl Lerner (film) [105]**

This is essentially the same as the book *Black Like Me*, but students often benefit from (and enjoy) seeing film adaptations of books they've read. If instruction time is an issue, it could be shown some evening outside of class and made optional.

***SuperVision: An Introduction to the Surveillance Society* by John Gilliom and Torin Monahan [69]**

This book is an introduction to surveillance studies, but it focuses heavily on issues of power and inequality and how surveillance promotes them. Surveillance is an increasingly important part of our society, and this book talks about how inequality is hidden and naturalized quite effectively (albeit in a limited context), as well as being short. For this reason it is effective for a general interest "Introduction to Identity and Oppression Studies" course, although of course it also works particularly well in the context of a Prison Studies curriculum.

***A Disability History of the United States* by Kim E. Nielsen [120]**

This book looks at the meanings disability has had in U.S. history, how they've evolved, and how this history related to racism, sexism, and other systems of oppression. It fits this course well for four main reasons: disability (esp. mental illness) is an important part of understanding prison populations; it's an unusual but important axis of oppression, and therefore helps students know how to examine other axes; it underscores the idea of a social construct because of its historical focus; and it integrates and sheds light on forms of oppression based of race, sex, ethnicity, and other things.

***Borderlands/La Frontera: The New Mestiza* by Gloria Anzaldùa [8]**

This book is an examination of multiple forms of oppression and ambiguous identities, named Borderlands because the author lives between different cultures geographically, linguistically, and in many other ways. The book presents a strong challenge to the idea of distinct and separate communities, emphasizing how all groups of communities have individuals on the "border" of two different communities (i.e., embodying elememts of both, but not fitting into either). As the book also discusses how different oppressed identities interact, this is an excellent book for examing intersectionality and understanding why it's important. The first half of the book is essay format but the second half is poetry, so the book also makes use of different ways of communicating and knowing than are standard in academics (and which work better for certain populations).

Recommended Reading: *Women's Voices, Feminist Visions: Classic and Contemporary Readings* by Susan M. Shaw and Janet Lee [146]

This is an anthology with a variety of historical and contemporary feminist readings. The chapters provide an overview on various topics, and then individual essays. While this should be supplemented by other essays and a few books that go more in depth, the anthology is strong and serves as an excellent basis for the class. It also functions as an ongoing reference guide for different gender issues and gives interested students a starting point for looking at different gender issues.

2.4 Introduction to Criminology

This course will introduce students to the basics of crime and the justice system, including how crimes are categorized, dominant theories on crime, how the justice system is structured, and the goals and functions of the justice system. The course will also examine some criminal justice policies and evaluate them in terms of criminological theories and statistics. It concludes by examining the theory of reintegrative shaming, both as a review of material from earlier in the course and as a theoretical basis for restorative justice.

Course Objectives

Students will be able to

1. categorize different types of crime.
2. describe different criminological theories of crime.
3. describe the goals and functions of the justice system.
4. critically evaluate criminal justice statistics and policy.
5. explain the theory of reintegrative shaming and its connection to other theories and data.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. explaining the categorizations, giving students the opportunity to categorize examples, and frequently using category labels in discussions.
2. presenting theories (or having students present them), explicitly comparing them, and analyzing different crimes in terms of those theories.
3. introducing goals, and then talk about policy in terms of goals.
4. analyzing policies in terms of different theories of crime and discuss statistics and outcomes.
5. teaching reintegrative shaming as a review of the other material.

Outline of Course Topics

Understanding Crime: The Measurement, Costs, and Dimensions of Crime This unit forms the basis of criminology and looks at the major methods and findings needed to understand crime and develop criminological theories and public policy. *Criminology* chapters 1-4 ('The Study of Crime', 'Measuring Crime', 'Crime and Its Costs', and 'Dimensions of Crime'). Homework mainly involves reading Conklin and a few supplementary articles, along with striaghtforward exercises in finding data and categorizing different crimes by type. (The purpose of these exercises is to familirize students with relevant databases (e.g., Bureau of Justice Statistics) and broad cateogries of crime as a foundation for future work, rather than to develop the skills in depth right now.)

Theories of Crime: Biopsychosocial Causes of Crime and the Development of Criminality This unit examines the major theories of criminology by looking at the causes of crime from a wide variety of perspectives, including the development of criminals and criminal organizations. Specific theories addressed include Social Control Theory, Differntial Association Theory, Labelling Theory, and Subcultural Theory. *Criminology* chapters 5-11 ('Biological and Psychological Explanations of Crime', 'Social, Cultural, and Economic Sources of Crime', 'Social Control and Commitment to the Law', 'Learning to Commit Crime', 'Opportunities and Facilitating Factors', 'Criminal Careers', and 'The Organization of Criminal Behavior'). Homework will focus on reading Conklin and supplementary articles, along with short-response pieces looking at different crimes in terms of different theories.

Responses to Crime: Community Responses and the Criminal Justice System The unit looks at how communities and the state address criminality, talking in depth about the philosophies underlying the criminal justice system. The community section talks about informal and community-organized controls, while the criminal justice section talks about the police, the court system, prisons, and the role of the victim in state processes. *Criminology* chapters 12-14 ('Community Reactions to Crime', 'The Criminal Justice System', and 'Deterrence, Incapacitation, Retribution, and Rehabilitation'). Homework will mainly involve reading Conklin, along with short-answer pieces relating these systems to theories of crime, and describing the criminal justice system in terms of its philosophies of punishment.

Public Policy and Crime: Assessing the Effectiveness of Public Policy This unit looks at different public policy responses to crime, including policies of different Presidents and various prevention measures, and analyzes them in terms of the different theories of crime and criminological data. *Conklin* chapter 15 ('Solving the Crime Problem'), heavily supplemented by other articles and readings. Aside from reading, homework will mainly involve summarizing different policies in terms of what they were, what theories they're based off of, and how effective they were (replace "were" with "are" for current policies). At the end of this unit students should choose a policy for the final project, which they will work on during the final unit.

Reintegrative Shaming: Integrating Theories of Criminology This unit talks about the theory of reintegrative shaming (which forms a basis for restorative justice practice) but serves as a review for the material in the rest of the course and a model for students' final projects. *Crime, Shame and Reintegration*, entire book. Homework involves reading Braithwaite and working on the final project, due at the end of the class. The final project is a paper that examines a public policy in depth, including the criminological theories that support it (and those that oppose it), how it relates to the criminal justice system, and an assessment of how effective it is at addressing crime. Because Braithwaite effectively integrates parts from all of the units of this class, it is hoped that the book will serve as a general model for how students can do so (while also being foundational for restorative justice work later in the curriculum).

Course Text(s)

Criminology by John E. Conklin [37]

This book, unlike many textbooks, focuses on concepts in criminology instead of merely looking at a variety of different crimes. As a result of this (and being written by a single author), the text is unified and fairly comprehensive, while still being quite readable. It addresses all the major topics in criminology, starting with definitions and measurements of crime, progressing to different theories of crime, examining community and legal reactions, and looking at history and other examples in the real world throughout. The book provides an excellent basis for a criminology course, and as an added bonus is quite inexpensive when bought used (in the second newest edition, from 2010).

Crime, Shame and Reintegration by John Braithwaite [26]

This book provides the basis of the criminological theory of reintegrative shaming, as well as many applications of the theory. It is a good book to end the course because it brings together many major theories of criminal behavior dominant in criminology, and it's an effective transition into the rest of the curriculum because reintegrative shaming is an important theory in restorative justice. It reiterates many of the major findings of criminology and makes policy suggestions, which students can review and judge using different evidence and theories in criminology.

Recommended reading: *Of Crimes and Punishments* by Cesare Beccaria [15]

This text was the founding work in penolgy, and is recommended primarily for its historical significance. However, much of the discussion is still relevant and insightful.

2.5 Restorative Justice Practices in the U.S. (prereq: Intro. to Conflict Resolution)

This course looks at the history, development, and practice of restorative justice. Specifically, it will building skills in analyzing different initiatives and deepen and focus skills in conflict resolution to practicing restorative justice.

Course Objectives

Students will be able to

1. explain the philosophy and development of restorative justice and its relation to retributive and community justice.
2. lead a restorative circle.
3. describe and analyze diferent restorative justice initiatives in the U.S.
4. further develop skills in reflective listening, assertion, communicating with themselves, and expressing appreciation.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. explaining the philosophy theoretically, by comparison with restributive justice, and by using examples; charting the development in terms of philosophy.
2. practicing circle processes in class, with students switching roles each time.
3. giving students examples of different RJ initiatives (and have them present if possible), then analyzing them in terms of theory and outcomes.
4. breaking down reflective listening, assertion, and conflict resolution into concrete, formulaic skills.
    - Reflective listening: students listen to other students/teacher and try to explain exactly what they said (i.e., the content of what they said, and the emotion behind it). In addition, for this and assertion students practice expressing and categorizing emotions so they can be more specific when they reflect feelings. (Extremely similar to Bolton from *Introduction to Conflict Resolution*; useful for review/practice.)
    - Assertion: teachers explain how to make three-part assertion messages using examples, then have the students practice translating other language into three-part assertions. (Note: the three part assertions here are slightly diferent from those in Bolton; students should now be able to do both types.) Teacher also models a method for making requests and has students practice it.
    - Communicating with Onself: students will practice translating internal dialogue into the language of NVC, and also note how they control their own feelings.

- Expressing appreciation: teachers gives students a framework for thinking about the goals of expressing appreciation, and then gives them a model for expressing appreciation and has them practice it.

Outline of Course Topics

Philosophy of Restorative Justice This unit looks at the theoretical and historical foundations of restorative justice. Following Zehr's *Changing Lenses*, it starts by looking at the current retributive system (focussing on victim and offender experience), examines different examples of justice systems in history, and talks about the beginning of the restorative justice movement. *Changing Lenses*, entire book, and *A Healing River* afterwards. Homework involves reading Zehr and writing short reflection pieces.

Circle Processes in Restorative Justice This unit looks at circles in restorative justice, focusing heavily on the mechanics of the circle processes and the values underlying circles. *Peacemaking Circles: From Crime to Community* (entire book), along with the film *Circles*. In-class work will heavily involve simulating restorative justice circles, and homework will mainly involve reading and reflecting on the circle process.

Assessing Restorative Justice Programs This unit looks at different assessment of restorative justice programs, starting with the *CoSA Evaluation Final Report* and then through *Victim Meets Offender*. Because both of these resources go more into depth about these models, in-class work will include some amount CoSA and mediation simulation. Homework will mainly involve reading and short analytic pieces about these models, with occasional reflection pieces.

Other Restorative Justice Initiatives This unit will examine other restorative justice initiatives, in particular making sure to include conferencing and reparative panels, in the context of *Nonviolent Communication*. In-class work will include discussions comparing different models and simulations of conferencing and panels. Homework will mainly involve readings and creative assignments where students choose and justify different restorative justice responses to different crimes/harms. The last few weeks be spent on simulating different students' proposed restorative justice solutions, and with the final assignment being a reflective piece on one of these simulations.

Course Text(s)

***Changing Lenses: A New Focus for Crime and Justice* by Howard Zehr [183]**

This book gives an overview of the current retributive system, the philosophical and historical basis of restorative justice, and a bit about RJ programs in the U.S. Zehr is a pioneer in restorative justice and as a theoretical justification this book is invaluable. It emphasizes the importance of paradigm shift but also describes the victim and offender experience in retributive justice in a way that is useful for someone trying to implement restorative justice in the current system.

Zehr's *The Little Book of Restorative Justice* is shorter and recommended as a reference book [185]; a version co-written with Ali Gohar is free online:

`http://www.unicef.org/tdad/littlebookrjpakaf.pdf`

***A Healing River* (film) by Cathie Douglas and Larry Moore [50]**

A Healing River takes a broad view of restorative justice, introducing how the process works and some of the major practitioners and theorists. The film also ncludes voices of victims and offenders, who speak to the impact of restorative justice on their own lives.

***Peacemaking Circles: From Crime to Community* by Kay Pranis, Mark Wedge, and Barry Stuart [129]**

This book explains the mechanics and theory of circle processes in restorative justice, with many real-world examples. Circle processes are a major part of restorative justice in connection to crime, as well as being useful for dealing with harms (or promoting communication) in general. The authors are all leading practitioners of RJ, and they relate their personal experiences to the text.

***Circles* (film) by The National Film Board of Canada [161]**

This film about Circle processes shows how sentencing circles work in a specific community. The intent is to give students a bit more depth with circles than they would get from just the Pranis book, as well as let them see actual circles being enacted. This film covers circles fairly broadly and would work as either an introduction or a conclusion to the cirlces unit (but probably best as a conclusion, since *A Healing River* is used to conclude the unit immediately before).

***CoSA Evaluation Final Report* by Kathryn J. Fox [61]**

This research report, produced for the Vermont Department of Corrections, evaluates the efficacy of Circles of Support and Accountability, a type of reintgration circle used in prisoner reentry. This report will both deepen students' understanding of CoSAs and provide a transition from the circles unit to the unit forcused more on evaluation. It is available online:

`http://www.doc.state.vt.us/about/reports/`
`circles-of-support-accountability-final-report/view`

***Victim Meets Offender: The Impact of Restorative Justice and Mediation* by Mark S. Umbreit [169]**

This book, while about 20 years old, provides a strong assessment of different restorative justice mediation program. It is a short book, and it effectively shows the process of how restorative justice programs can be evaluated, while also giving a lot of evaluation of mediation as a process generally. While the focus of the book is assessment of programs, it details a number of early models of restorative justice mediation in the U.S. (including examples of mediations), and even talks about what potential mediations programs should or must make sure to do.

***Nonviolent Communication: A Language of Love* by Marshall B. Rosenberg [137]**

This book describes a general method for communication, in terms of both listening and expressing onself clearly. In that sense, it is useful for review, since it is similar to *People Skills* and is at the end of the restorative justice sequence, although since it takes a different approach in some parts it will deepend students' understanding and abilities to communicate. One thing this book

does particularly well is talk about self-communication, which is important for people to maintain balance in their lives generally (and resolve conflict). Finally, this book is chosen because it is a short, easy read; this allows the end of the course to examine models like restorative panels and tie together any loose ends.

2.6 Research and Information Literacy (prereq: Statistics for Social Sciences)

This may already be covered by another department (or as a general ed. requirement): if not, it is designed to be useful to people from other departments. Students will have research projects of varying length, including research in other classes, and develop their information literacy through those projects along with theory readings.

Course Objectives

Students will be able to

1. determine the nature and extent of the information needed.
2. use information effectively to accomplish a specific purpose.
3. access needed information effectively and efficiently.
4. design, administer, and analyze surveys.
5. write a successful reseach proposal in complicance with the school's Institutional Review Board.
6. evaluate information and its sources critically.
7. write effective research papers, and revise them effectively with feedback.
8. write an annotated bibliography.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. giving students different research questions, having them determine what data are needed, and discussing it as a class.
2. giving students tasks that they need certain information to accomplish.
3. showing them how to search in major databases and discussing (and practicing) different strategies for researching.
4. having students design survey questions to find different types of information, and practice surveying people.
5. having someone from the IRB talk about their policies in class and discussing ehtical research guidelines.
6. *For evaluating sources*: showing them what to check for when evaluating a source (author, citations, journal of pulication, etc.) and having them write evaluations of different sources.

   For evaluating information: teaching students the structure of an argument and having them practice breaking down arguments into their parts and analyzing them.

7. having students write a research paper during the process of this class (the paper can be from another course but must be workshopped in this one).

8. showing students how to make an annotated bibliography (and what it is) and having them write one (with a paper for another class if possible).

Outline of Course Topics

Defining a Topic and Research Question This unit focuses on the beginning of the research process. *The Craft of Research* chapters 1-4 (section "Research, Rsearchers, and Readers" and chapters "From Topics to Questions" and "From Questions to a Problem"). Homework mainly involves reading and short written assignments of looking at different situations and determining what information is needed to solve given problems.

Finding and Assessing Sources This unit focuses on finding effective sources (and relatedly, evaluating sources). *The Craft of Rsearch* chapters 5-6 ("From Problems to Sources" and "Engaging Sources"), along with supplementary articles about research. This unit also includes a discussion of (and practice in) survey design. Both in-class work and homework mainly involves searching and evaluating sources (though homework also involves reading); as much as possible, the course attempts to use the questions from the previous unit to guide its searches. In addition, students will learn how to write a research proposal for an Institutional Review Board during this unit to address one of the questions.

Building Arguments and Evaluating Information This unit focuses on the structure of effective arguments and more general skills in evaluating information. *The Craft of Research* chapters 7-11 (section "Making a Claim and Supporting It"), along with supplementary articles and/or another short text. Homework will mainly involve reading and outlining arguments, along with writing a short essay answering one of the original research questions.

Writing a Research Paper (Longest Unit) This final unit involves actually writing and redrafting a research paper. *The Craft of Reearch* chapters 12-17 (section "Planning, Drfting, and Revising"). Homework involves expanding, revising, and redrafting the essay from the previous section. Students may instead use a research paper from another class (and if they wish, they may bring in multiple papers from other classes for feedback). Class mainly involves short writing exercises and editing each other's papers. The final project is a research paper, from this class or another one.

Course Text(s)

***The Craft of Research* by Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams [23]**

This book examines the entire process of doing research, from deciding a topic and research question to revise and creating a final product. The text is chosen mainly because it is clear, readable, and fairly short; it also is noteworthy for its specific discussion of components of an argument and its heavy focus on revising and communicating clearly with a final product.

The text is free online:

http://www.andrewwaugh.com/courses/pols455/craftofresearch.pdf

2.7 Poverty in the U.S. (prereq: Intro. to Identity and Oppression Studies)

This course will examine poverty in the U.S., focusing on how it impacts people, families, and communities. It will look at different theories of what causes poverty and particularly focus on links between poverty and crime. It also attempts to capture and convey the experience of being poor in the U.S. and how that affects available institutions in a community and daily life for the poor.

Course Objectives

Students will be able to

1. point to effects of poverty at the individual, family, and community level and explain how they interact with race, gender, and disability.

2. describe and analyze different methods of measuring poverty in the U.S. and how poverty can be "invisible".

3. describe how inequality intereacts with a number of social issues at a societal level, and how it affects the poor in particular.

4. propose and justify ways to combat poverty in the U.S. and talk about these in terms of the history of movements to fight poverty.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. discussing poverty in terms of these three levels and in terms of intersectionality throughout the class.

2. showing the students different methods of measuring poverty and what aspects they (fail to) measure well.

3. looking at the broad effects of inequality and examine various social issues in more depth for poor populations.

4. talking about their own expeiences fighting poverty and contextualize others' experiences that we read about.

Outline of Course Topics

Portrait of Poverty This unit is dedicated to helping students understand the experience of poverty at individual, family, and community levels. It will specifically address through supplementary articles the deinstitutionalization movement and challenges in mental health care, connecting these to general infrastructural problems. *Amazing Grace*, along with supplementary articles about mental health challenges. Homework mainly involves reading and short pieces reflection pieces.

Invisibility and Prevalence of Poverty This unit transitions to looking at poverty in a more systemic way, focusing on before and during the War on Poverty and comparing them to present day. *The Other America*, articles about the Great Society programs, and supplementary articles about poverty more recently. Homework mainly involves reading and a short paper analyzing poverty systemically.

Inequality and Society This unit follows Pickett and Wilkerson's text and examines the relationship between a number of social problems and economic inequality in developed societies. *The Spirit Level* and short supplementary articles for each class. Homework mainly involves reading, as well as short responses to two of the chapters.

Ending Poverty This final unit focuses on addressing and combating poverty. It integrates diverse types of information, mainly from Baptist and Rehmann. *Pedagogy of the Poor* and supplementary articles. Homework mainly involves reading and working on the final project, which is a proposal for addressing poverty through community work.

Course Text(s)

***Amazing Grace* by Jonathan Kozol [99]**

This book (like everything Kozol writes) provides an in-depth, personal protrait of poverty in the U.S., in particular focusing on its relationship to segregation. Kozol looks at individuals, families, and communities and how poverty and inequality affect them in terms of health, violence, employment, life opportunities, religion, and hope. The book is emotionally powerful and paints the effects of systemic trends on people's lives.

***The Other America* by Michael Harrington [80]**

This was an extremely influential study of poverty in 1962, arguably responsible for the War on Poverty and directly providing the ideas of Medicare, Medicaid, food stamps, and other social programs to help the poor. The book talks about how poverty continues to affect people in an otherwise prosperous nation, and additionally notes how the structure of society works to conceal this poverty from mainstream society. Despite being 50 years old, it still provides a useful portrait of the experience of poverty in the U.S.

***Pedagogy of the Poor: Building the Movement to End Poverty* by Willie Baptist and Jan Rehmann [13]**

This book, co-written by a poverty activist and and academic theorist, combines many different types of knowledge about poverty including personal experience (in the form of an interview), scholarly theory, case studies, in-class debates, and both authors' reflections on poverty. It is useful because it addresses poverty from many different perspectives, including talking about social activism to end poverty and some of the struggles involved. It also discusses a bit of the history of poverty in the U.S.

The Spirit Level: Why Greater Equality Makes Societies Stronger by Kate Pickett and Richard Wilkinson [177]

Taking a global perspective, this book examines the effects of inequality *as distinct from the effects of poverty*, i.e., why poor people are more harmed in a rich than poor society. The book looks how how inequality is related to a wide variety of issues, both across countries and across U.S. states; it therefore provides an excellent jumping-off point for discussing how those issues affect the poor. This book is fairly long but quite readable.

Recommended Reading: *Poor Economics* by Abhijit Banerjee and Esther Duflo [12]

This books looks at global efforts to combat poverty, with a focus on understanding the lives of the poor and why some interventions fail (as well as how to fix them). The book is not included in the course mainly because this course is U.S.-focused, but it is an excellent read in general and effectively combines some of the strengths of Kozol and Baptist/Rehmann.

2.8 History of Prisons (prereq: Intro. to Criminology)

This course looks into the appearance and development of the modern prison, focusing on reforms and guiding philosophies and how they have changed over time. It contexualizes the development of prisons in terms of other institutions like schools, mental hospitals, and systems of racial control.

Course Objectives

Students will be able to

1. explain the appearance and development of modern prisons, along with a number of reforms in prison history.
2. chart the development of prisons in the South and how this influenced their development across the country.
3. relate the prison boom to the previous Jim Crow system and detail the mechanics of how the prison system oppresses poor blacks and Latinos.
4. describe the experience of life inside a prison and how individuals' experiences relate to systemic trends.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. telling the history or prisons in terms of the theoretical goals of the justice system and how they changed over time.
2. examining the history of prisons in the South in-depth and constantly relating it to broader carceral trends.
3. talking about the mechanics and tying felony-based restrictions directly to Jim Crow restrictions.
4. examining Santos' experiences in depth and tying them to other things we've been studying in the course.

Outline of Course Topics

Prison Origins, Reform, and Abolition This unit is an introduction to prison history and briefly covers the origin of modern prisons, looks at some reforms in prison history, and examines the prison abolition movement. Students will read *Are Prisons Obsolete?*, along with supplementary articles. Homework mainly involves reading and reflection pieces.

Prison History of the South This unit looks at how prisons have deveoped in the Southern United States and relates it to developments in the prison system generally. *Texas Tough*, entire book. Because this book is fairly long and dense there w ill not be a lot of other homework, and classes will mainly involve discussion of material and tying it to broader historical trends.

The Prison Boom This unit examines the prison boom, from roughly 1972 to the present. Following Alexander it focuses heavily on race and the mechanics of oppression, while examining what led to the prison boom and what this meant for the philosophy of punishment. *The New Jim Crow*, along with supplementary articles. Homework mainly involves reading and short essay pieces.

Life Inside of Prisons This final unit looks at the experience of life inside a prison, making attempts to connect this to prison history and the prison boom. *Inside* and supplementary articles. Homework mainly involves reading and working on the final project, which describes some important part of prison life in terms of its history and wider significance.

Course Text(s)

***Are Prisons Obsolete?* by Angela Davis [45]**

This very short book gives a (fairly detailed) overview of the history of prisons in terms of race, reform, gender, and connections to the private sector. It also makes the philosophical case for the prison abolition movement, which is relevant to many of the students in this major field and valuable to consider in the context of restorative justice.

The book is available as a free pdf online:

`http://www.feministes-radicales.org/wp-content/uploads/2010/11/Angela-Davis-Are_Prisons_Obsolete.pdf`

***The New Jim Crow: Mass Incarceration in the Age of Colorblindness* by Michelle Alexande [3]**

As the title implies, this book details how anti-black racism has remained a major force in the U.S., specifically how the prison system functions as the primary institution reinforcing racial inequality in the U.S. today. Drawing parallels between the denial of rights under Jim Crow and restrictions that prisoners and ex-felons face, this book further elaborates how our current system grew directly out of those systems. The book addresses how "colorblind" policies lead to the continued oppression of blacks in the U.S. It details the mechanics of how the current system works, why it disproportionately affects poor, black (and brown) males, and why the prison system is an integral part of society in the U.S. currently.

***Texas Tough: The Rise of America's Prison Empire* by Robert Perkinson [124]**

This book gives a detailed history of prisons in Texas and the rest of the South, detailing penal philosophy, connections to slavery and racism, and attempts at reform. It connects the development of prisons in Texas with how they have developed in the U.S. generally, filling in important gaps in most prison histories (which focus on the Pennsylvania and Auburn models) and explaining how the U.S. became the prison capital of the world. While it focuses most on Texas and the south, it covers leading prison reform efforts throughout the country from the beginning, and by the end is able to use the Texas system to describe all the essentials of the U.S. prison boom, including the rise of the security state domestically and abroad following 9/11. The book is extremely detailed and uses a wide variety of primary sources to support its argument. It is long and fairly dense at times, but this is made up for by the level of detail.

***Inside: Life Behind Bars in America* by Michael G. Santos [142]**

This book, written by the only person in U.S. history to experience every security level of the U.S. prison system, documents life inside prisons in the U.S. It spans prisons across the country and includes stories from other prisoners Santos met. While the work is primarily anecdotal, it shows a lot of systemic problems in the prison system and even talks about some of prison history. In addition, the work shows how prisons are violent and anti-rehabilitative for many people while also showing how for some people (the author, and one of his friends) they can be genuinely rehabilitative (almost in spite of themselves). True to prison dialects and informal systems, this work is unique in its comprehensive, personal look at prisons from an inside view, by an uncommonly skilled writer (especially given the low levels of education characterizing most prisoners).

Recommended Reading

***Discipline and Punish: The Birth of the Prison* by Michel Foucault [60]**

Discipline and Punish is one of the most famous books about prisons, and one of the most cited books in the humanities in general. It chronicles the history of punishment and how prisons developed in the U.S. and elsewhere, connecting all of this with discipline and articulating in detail the development of disciplinary techniques. Foucault is strongest in his description of disciplinary techniques in this book, but unfortunately a bit weak in terms of describing actual prison history.[7] However, this book connects prisons to other social institutions in a way that more than makes up for this weakness.

***The Culture of Control: Crime and Social Order in Contemporary Society* by David Garland [65]**

This book details major developments in the penal system from 1970-2000, the period of the prison boom, in both the U.S. and England. Focusing on both the technical aspects and the guiding philosophies, it is a useful work for understanding this period. Importantly, Garland makes sure to note that the history was not predetermined and could have gone differently, so it is not read as inevitable. As a guide to understanding the U.S. during the prison boom this work is extremely useful, and the comparison to England is useful for perspective.

***Instead of Prisons: A Handbook for Abolitionists* edited by Mark Morris [97]**

This book from the 1970s details some of the abuses of the prison system and looks at various alternatives to it. While it does not say as much about prison history as I would like, it is a book I feel students should potentially read for this major and am not sure where to put it. It doesn't seem to fit restorative justice particularly well either, and since prison abolition was an important theme in the 70s and reveals a lot about how the system was, this book might be useful for serving as a comparison to prisons today.

[7] See, for example, David Garland's critique in *Punishment and Modern Society* [64, p. 157-162]

2.9 Electives: choose two

These electives allow students to pursue particular areas of personal interest within the field, which helps them integrate knowledge overall, prepare for their community work, and become committed scholars in the field. Classes have individual content goals, as well as universal elective goals, below. Note that students are required to take at least two electives and required to take a course relating to their internship; this may be one of the two electives or it may be a course from another department. These courses are intended as recommendations or examples more than requirements; an individual school should do whatever makes sense given its teachers' strengths and interests and local community organizations.

Course Objectives

Students will be able to (universal goals)

1. integrate knowledge through sbjects of personal interest.
2. develop theory that will be useful in community volunteer work.
3. find, read, and present related books of personal interest.

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. making these courses electives with wide-ranging (and interesting!) focuses.
2. having the themes cover major areas that students could do community work with; when students have specific ideas about where they plan to volunteer, addressing issues that would be directly relevant.
3. assigning students the task of finding two pertinent books (in each course), reading them independently, and then presenting about them and tying them into other themes in class. (If the class is small enough, having the students lead discussions about them; alternately, instead of reading them independently, they can form groups that read the same book, present on it, and then lead a discussion together.)

2.9.1 Quantitative Analysis (prereq: Research and Information Literacy)

This course focuses on analyzing and gathering statistics. Essentially, it will expand students' strength in using the regression model and other statistical analyses, as well as their ability to create and analyze quantitative data visualizations. It will include analyzing information from major government databases (e.g., Census data).

Possible course texts: *Econometrics* by Woolridge [179]; *The Functional Art* by Cairo [30]

2.9.2 Qualitative Research Methods (prereq: Research and Information Literacy)

This course looks at different methods for qualitative research, as well as how to analyze qualitative information. Specific skills covered include survey design, administration, and analysis; interview technique, coding, and analysis; and ethnography writing.

2.9.3 Prisoner Reentry (prereq: History of Prisons)

This course examines probation, parole, furlough, and unsupervised release. Specific focus is on the challenges prisoners face in reentry and how to support them in this.

Possible course texts: *But They All Come Back* by Jeremy Travis [166]; *When Prisoners Come Home* by Petersilia [125]

2.9.4 Causes of Violence and Crime (prereq: Intro.'s to Conflict Resolution, Criminology, and Identity and Oppression Studies)

This course looks at what causes people to act violently and commit crimes, with a focus on violence. It integrates and extends understanding of conflict, theories of crime, and identity/oppression.

Possible course texts: *Violence* by Gilligan [68]; *The Creation of Dangerous Violent Criminals* by Athens [10]

2.9.5 Identity and Oppression Studies Seminar (prereq: Intro. to Identity and Oppression Studies; History of Poverty highly recommended as prereq/coreq)

This course is driven by student interest and needs for their community work (i.e., it focuses on sex/gender if they want to work at a women's center, etc.). The structure is highly dependent on what students want to study, with no regular course texts.

This course examines nonviolent resistence, both historically and in terms of the techniques and mechanics of nonviolent resistence. This will include a discussion of political prisoners, as well as negotiation with violent oppressors.

2.10 Community Work and Reflection Class (pre/co-req: Research and Information Literacy and appropriate elective)

This class involves volunteering in the community and reflecting on it in an academic setting that connects it theory. While this is a semester-long class, students are required to do a year of community work (although the other semester could be a different organization).

Course Objectives

Students will be able to

1. contribute to their communities in a positive way
2. apply academic theory to real life
3. integrate actual experiences into the building of theory
4. produce original research
5. perform First Aid/CPR

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. helping students find places to volunteer or do community work; organizing transportation when necessary; estbishing partnerships with local organizations
2. helping students explain their experiences by using theory, and providing them with relevant theory when necessary
3. having students reflect on how well theory fits with their experiences; having students decide what is important in their experience and should be the focus of theory
4. choosing a research question at the beginning of the term; helping students with data issues; forming partnerships with the local organizations (who may have their own research questions students can use or contribute to)
5. training and certifying students in First Aid/CPR during the first weeks of this class

Course Text(s)

Outline of Course Topics

First Aid/CPR Training This unit will take up the first two weeks, when students are also deciding what to examine at the organization where they're volunteering. In-class work mainly involves First Aid/CPR training, while homework involves deciding on a research question. At the end of this unit, students will be First Aid/CPR certified[8].

[8]First Aid and CPR are included in this curriculum because they're fairly basic, but potentially life-saving, skills, and therefore something everyone should have to contribute to their communities. Possibly it makes more structural sense to run the training outside of the course, but this way keeps student workload in check.

Community Volunteering The rest of the class will have a much looser structure. Students will have weekly reflections about their community work (and time in class will also be primarily be spent reflecting, as well as presenting), readings co-determined by the student and instructor, and two reflection papers (one halfway through the semester, one at the end). Students and professor will attempt to integrate past academic work as much as possible so students challenge, reinforce, and deepen their understanding of what they learned previously. Students are also expected to take an active interest in each others' projects; some of the weekly reflections may be replaced by writing about *other* student's experinces described in class.

2.11 Senior Thesis (prereq: be a senior in this major field!)

This course must fulfill whatever the college's normal thesis requirements are, with the expectation (although not requirement) of tying in community work to the thesis as well.

Course Objectives

Students will be able to

1. synthesize knowledge and show what they've learned
2. examine a specific issue in more depth than any previous academic work
3. graduate college!

Teaching Methods

Teachers will give students the tools to accomplish these goals by

1. discussing the students' thesis topics in a seminar format, reviewing and tying together knowledge (and skills) from the common knowledge base
2. devoting an entire class to examing the issue of interest, which already has support from previous academic (and ideally, community) work as a starting point
3. serving as advisors, teaching time-management and related skills

Course Text(s)

Student- and thesis-dependent; no general texts

Outline of Course Topics

No formal structure for the course here. Students are expected to make progress on their thesis each week, with adequate progress co-determined by students and professor. Work in class will mainly involve students giving each other feedback or the class doing writing and research activities/exercises. Final project: senior thesis!

Part II

Public Policy and Evaluation

Chapter 3

Prison Statistics, Invisibility, and Inequality: How Excluding Prisoners from Statistical Measures Conceals Social Problems and Continuing Inequality

List of Figures

3.1 Invisible Inequality

This paper examines how prisoners are excluded from or miscounted in major statistical surveys, and what effect this has on perceived inequality in the U.S. I start by looking at the example of unemployment measured on the Current Population Survey (which excludes prisoners), and by reporting and contextualizing analyses from scholars Bruce Western and Becky Pettit that account for prison and jail populations. After discussing issues involving the use and measurement of statistics in social sciences generally, I return to unemployment as an example and extend Western and Pettit's analysis by another decade (with some modifications). I then turn to other measurements derived from the Current Population Survey, specifically rates of high-school dropouts, wages, and voting rates, explaining and commenting on Western and Pettit's analyses in these areas. I then examine the Census (which measures prisoners differently) and report the analyses from Prison Policy Initiative, illustrating the importance by using Brown County and Cook County from Illinois as examples. I briefly give an example of statistical mismeasurement in another area, health, and then conclude with another discussion of measurement in statistics.

Many scholars, in particular Angela Davis, have talked about how prisons hide social problems from public view by locking up visible perpetrators (and victims) of these problems instead of addressing the root causes of these problems. In this way prisons "disappear" a variety of social prolems: "[h]omelessness, unemployment, drug addiction, mental illness, and illiteracy are only a few of the problems that disappear from public view when the human beings contending with them are relegated to cages." [44]. The fact that they are sited far away from population centers functions to hide social problems geographically as well [131]. This paper examines how prisoners and hidden statistically. More specifically, it looks at the way many national surveys exclude prison populations from their samples, thereby misrepresenting many social problems. Misrepresentative

data makes it more difficult to make good public policy, and also makes evaluation much more difficult. Incarceration is disproportionately concentrated among poor, minority, uneducated males, so to the extent that statistics deal with those populations they will generally be more distorted. In particular, statistics that measure racial, class, and educational equality will be overly optimistic. This paper examines some social problems where this effect is especially significant, explains how to correct for it mathematically when possible, and discusses the broader significance of this for social science research.

3.2 Unemployment

Understanding unemployment is essential to understanding U.S. life because employment directly affects people's livelihoods at a personal level and economic productivity at a societal level. As a result, the measured unemployment statistic is one of the most important economic indicators and significantly affects elections and public policy [128]. In recent years this is most clearly seen in Obama's stimulus package (a very expensive and influential public policy designed to directly combat unemployment), as well as his reelection campaign (in which Romney blamed him for persistent high unemployment). As an economic indicator, unemployment can show how well an economy is maximizing its potential, as well as individuals' ability to find work and thereby be able to provide for their families. In part because it is so widely used and so important, the distortions resulting from unemployment are probably the most widely understood: economists regularly use the nonemployment rate alongside the unemployment rate as a measure of economic health, and newspaper articles regularly discuss the limitations of unemployment as a statistic.[1]

These two measures are different because unemployment does not directly measure the number (or percentage) of people not employed, but rather the portion of the labor force that is unemployed. For these purposes the "labor force" is defined as the group of people who are employed or have sought employment in the past six months [128]. In a lot of ways this makes perfect sense, because nonemployment among groups like children and retirees is expected and not a worrying sign for the economy. It also excludes people who have given up on job-seeking because of a poor economy, which definitely limits its usefulness during recessions. Most relevantly here, it excludes people in the military, mental institutions, and prisons. Excluding each of these groups has its own set of advantages and disadvantages, but excluding prisoners is particularly problematic because prisoners would normally prefer employment to prison and our economy would be more productive if prisoners were working (and not committing crime). To the extent that large numbers of people cannot work because they are in prison, our economy is less productive and our statistics fail to measure that. As with most statistics that exclude prisoners, the effect will be particularly pronounced when looking at populations that are poor, lacking in education, racial minorities, young, and males, because these groups are disporportionately imprisoned.

Comparing nonemployment statistics that include and exclude prisoners, we find that the measured nonemployment is about 1 percentage point lower than it would be if it included prisoners. This is important in itself, and becomes even more significant when we consider that imprisonment is far from evenly distributed: some neighborhoods have many more residents in prison, largely as a result of how disproportionately minority men are incarcerated (although there would be significant variation anyway) [116]. For certain groups, these statistical differences can be quite dramatic.

[1] There have been examples of this from CNN, Huffington Post, NBC, Forbes, and the Wall Street Journal in the past three years, among other places [18, 35, 41, 55, 94].

Young black men (ages 18-34) without a high school education have a roughly 26% chance of being employed, a 37% chance of being in prison, and a 36% chance of being neither employed nor imprisoned (as of 2008) [127]. If we assume that almost all of the people in this age group who are not in prison or employed are unemployed (a reasonable assumptions, since they are old enough to work and too young to retire, and relatively few people are in the military or mental institutions), we can calculate the rate of unemployment by dividing 0.36 (the rate of unemployment) by (0.26 + 0.36) (the rate of labor force participation). By this calculation the unemployment rate of black high school dropouts aged 18-34 is around 0.58. The rate of non-employment, which includes the unemployed, prisoners, and anyone else who is ot working, is substantially higher, at 0.74. Again, these statistics are country-wide: at local levels, or looking at smaller groups, the differences in some areas are even more dramatic (conversely, the differences in some areas are less dramatic). In states like Louisiana, which has an especially hgh imprisonment rate, the difference betwee the unemployment rate and the nonemployment rate is likely to be higher, which means lack of employment in this subpopulation is markedly understated.

We know that prison-related distortions in the unemployment rate will be significant for specific population groups, specifically the groups that have a higher proportion of people in prison. Because blacks are imprisoned at a much higher rate than whites, the effect on the employment rate for blacks is much higher [116]. This difference means that statistics looking at the difference between black and white employment rates will also be skewed. We can calculate statistics for the white non-employment rate using the same method as above, and then calculate the racial difference in employment by dividing the black rate by the white rate. Because the prison population has expanded significantly and continually since about 1970 [164, p. 1], the effect has grown larger over time. Figure 3.1, a graph of the employment ratio in different years, shows how this statistic became increasingly distorted by the prison boom. The bottom curve displays the black-white jobless ratio as normally measured, while the top curve shows the ratio when prisoners are included; both curves have been smoothed to illustrate the general trend better. For clarity, a black-white jobless ratio of 2 on the graph below (as in 1980 when inmates are included) means that there are twice as many jobless blacks as jobless whites (both noncollege, ages 22-30).

In summary, unemployment understates the rate of involuntary non-employment of the general population, and does so dramatically for certain disadvantaged groups. This effect has also increased over time, so that by 2000 icluding prisoners increases the black-white nonemployment disparity ratio by about 13% for noncollege males aged 22-30. While the methodology used to calculate the unemployment rate makes sense for many purposes (in particular, for measuring the bargaining power of workers relative to that of employers), it has some really important side effects when the goal is to understand economic health more generally. To really understand the U.S. labor force, one must adjust the statistics to include prisoners.

3.3 Statistics in the Social Sciences

Before examining many of the statistics that have been distorted by imprisonment in this way, we should consider some important statistical issues involved. First, when analyzing something statistically it is essential to decide what type of statistic you want ahead of time and then search for that, rather than deciding to use whichever statistics leads to the strongest result. In unemployment this issue is complicated because prisoners, the military, and people in mental institutions could all be reasonably included or excluded from the statistic. Despite the existence of prison labor,

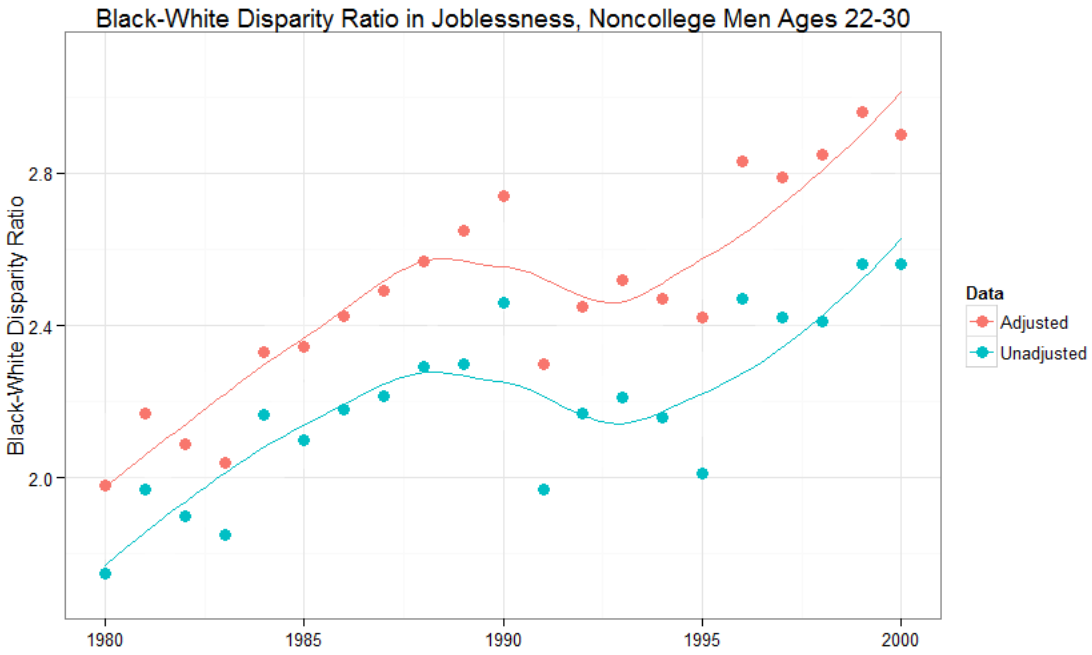


Figure 3.1: This graph was created by extracting data from a graph by Bruce Western (details about this process are in the methodological appendix) [176, p. 93].

I chose to count prisoners as jobless because it would be economically healthier for them to have jobs outside of prison if possible, and also because many of them would likely be unemployed if not in prison (see "Wages", below). The military and people in mental institutions compose a comparatively small part of the population, and I excluded them both because of a lack of good data and because the effect would likely be relatively small next to the effect of including prisoners. (If I were to include them, I would count the military as employed and mental patients as unemployed, which means they would cancel each other out to some extent.) In some statistics later, like the high school dropout rate, deciding how to count people from special populations like this may be straightforward, because I can just pool the data and recalculate it including prisoners. Methods for couting other statistics, like wages, require more of a argument, because there are many reasonable alternatives. The reason it matters to do this ahead of time is that if someone looks at many possible statistics before choosing which to use, that person may unconsciously choose whichever one seems most impressive (or be left open to the accusation of doing so) and by choosing that method will exaggerate the effect (in terms of both practical and statistical significance).

The second major point to consider is that many different populations can be excluded as a result of how statistics are sampled. Broadly speaking this happens in two ways: surveys are household-based, meaning households are selected randomly and surveyed, or they are based on the eligible population. Unemployment statistics, which come from the Current Population Survey, are actually both; they come from a household survey *and* look only at the labor force. Some other major federal surveys that are household-based includes the National Health Interview Survey, the National Survey of Drug Use and Health, and the National Crime Victimization Survey. The

Census is an exception to this that counts prisoners in a fairly unique way, which is also important and will be discussed later in this paper. Often we can adjust the statistics with respect to prisoners to get the information we actually want, but other population are more difficult to account for. People who are homeless (or lack stable housing) will be excluded from household surveys (as well as mental patients and the military), while other groups can be excluded if they are not part of an eligible population (e.g., permanently disenfranchised felons are excluded from voting statistics). I focus on prisoners in this paper in part because they are a large and easily measurable population, which means that statistical effect is likely to be relatively large, and it is possile to adjust for. But these statistics still fail to tell the whole story, or even represent the populations they should measure.

In discussing the weaknesses of statistics here, it is important to remember that all of the methodological choices in these statistics are intentional, either because they intentionally exclude noneligible populations or because household surveying is an effective way to get a relatively random sample, and the unintentional distortions caused by not surveying certain groups are assumed to be small. The unemployment rate is used by economic theorists to look at how markets will work, because what matters there is the relative power of workers and employers, and that is determined to a significant extent by how much of the labor force is employed. Since these statistics are gathered for the benefit of economists, it is totally logical to exclude these populations, because that makes the statistic much more useful for their purposes. When it becomes a problem is when the statistic is used for something different from what it was originally gathered for. When unemployment is used as a measure of economic well-being, as it continually is in politics, particularly elections, many of these methodological choices are counterproductive, because they hide rather than reveal economic problems. Any time a statistic is being used for something substantially different than what it was designed for, it is important to see whether the methodology used to calculate it has any important, negative effects on the result.

As a final note, I want to emphasize that when I advocate recalculating a statistic or using another one (e.g. using non-emplyoment rather than unemployment rate), it is not because the other statistic is bad or useless, or even because it is wrong per se. Even though it excludes prisoners, the unemployment rate is useful for a number of reasons, including measuring how productive our economy is. In most cases it makes sense to exclude retired people from the unemployment rate, because having an economy where a lot of people are retired probably reflects to some extent conscious choice more than low economic productivity (or might reflect a confounding variable like age). Again, the methodology of how unemployment is measured is highly intentional, and appropriate for many purposes. Only recently, as prisons have become a common life experience for many populations and are no longer always negligible in their statistical effects, does it make sense to change how statistics are measured with respect to prisoners. While I may use the language of "correcting" statistics for clarity and ease, the process is not so much "correcting" (which implies the first statistic is wrong) as it is "improving". I argue here that for most purposes, the modifications I make result in more useful statistics, but it is important to remember both that the unadjusted statistics are still useful, and that for some specific uses the unadjusted statistics are superior.

3.4 Unemployment, Reconsidered

Figure 3.1 from the previous analysis of joblessness is based on Bruce Western's work, which extends from 1980-2000. Using later versions of the Current Population survey and reports from the Bureau

of Justice Statistics I was able to extend this analysis another decade through 2011 (a complete description of my analysis, including the R code used, can be found in the appendix to this chapter). My version is somewhat simplified in that I adjust only for prisoners (not jail inmates or military personnel, which Western includes), and my numbers from the Current Population Survey data are slightly different even when unadjusted. As a check for validity I compared my calculated rates of the unadjusted data from 1980-2000, and for adjusted data from 1996-2000 (I was unable to replicate the adjustment before 1996). My numbers are slightly different but otherwise the trend is the same for unadjusted years, and the difference caused by adjustment is comparable (though a bit lower) for the adjusted years also. Figure 3.2 compares my data with Western's, with my data in red and his in green.

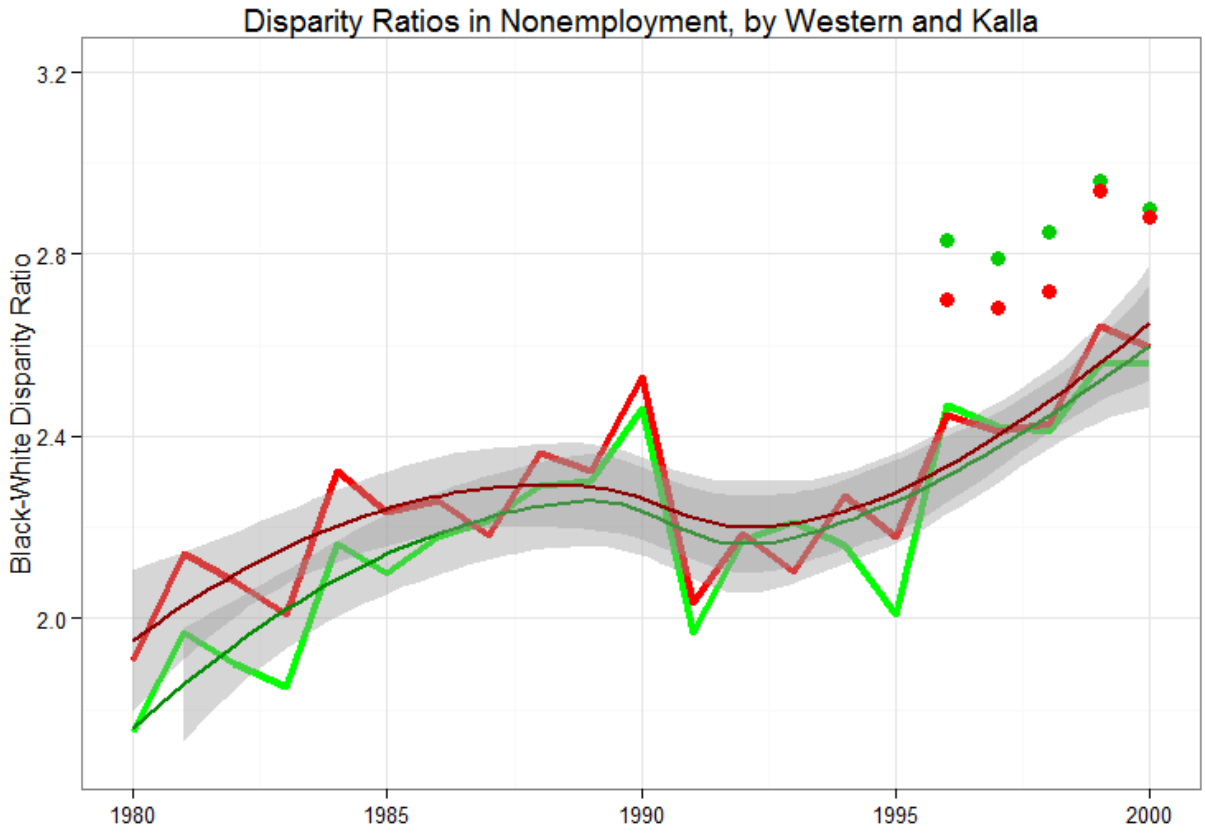


Figure 3.2: This graph compares my analysis of CPS data with Western's; my data are in red while his are in green. For unadjusted data, the spiked line compares data points directly while the smooth lines show the trend. The points in the upper right corner show our calculated adjusted disparity ratios. My data match fairly closely with Western's, especially toward the end, but my adjustments are only about 3/4 as big because of my more conservative methodology.

In my analysis, we find that the inequality of joblessness remains roughly constant from 2000-2007 before dipping sharply in 2008 and reaching record lows the following three years (this is true for both unadjusted rates and adjusted rates). Figure 3.3 shows the unadjusted rates of joblessness

from 1979-2013 and the adjusted rates from 1996-2011.

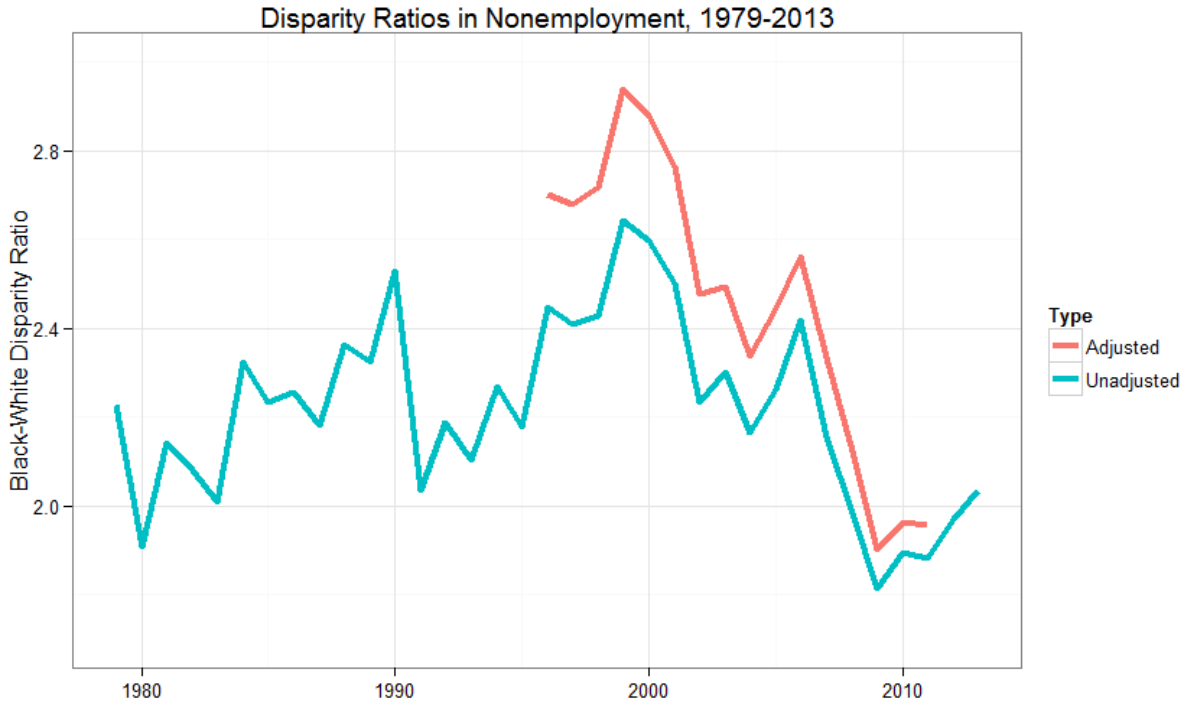


Figure 3.3: This graph contains the unadjusted disparity ratios for years 1979-2013, and the adjusted ratios for 1996-2011.

Note that the disparity ratio in nonemployment declined from about 2006 to 2009, reaching an all-time low in 2009. If we look more closely at the data used to calculate inequality in joblessness, detailed in Figure 3.4, it is clear that this does not represent a victory for young black men. In fact, the absolute difference in rates of joblessness remains constant or even increases during this period. The major result here is that rates of joblessness for both groups rose rapidly in 2008 because of the economic recession. While the level of joblessness increased more in absolute terms than the level of white joblessness, the disparity in change in jobless was lower than the original disparity, so that overall the disparity in joblessness decreased. This graph shows the unadjusted rates of joblessness for young noncollege blacks and whites because I have a fuller data set and can show the trend across time more effectively. The unadjusted rates would be higher for both races, so Figure 3.4 understates nonemplyment levels somewhat.

As always, none of these graphs tells the whole story, or even the best story. Though I include prisoners, my adjustments exclude jail inmates and military personnel from consideration, and both mine and Western's exclude homeless and other hard-to-reach populations (which would likely increase the rates of joblessness, though the effect on the disparity ratio could be positive, negative, negligible, or variable depending on the year). The disparity ratio tells more about inequality but obscures the massive increase in joblessness for young noncollege black men in the 2008 recession. The best way to understand the situation, therefore, is to look at multiple graphs, such as Figures 3.3 and 3.4, and understand what each is saying.

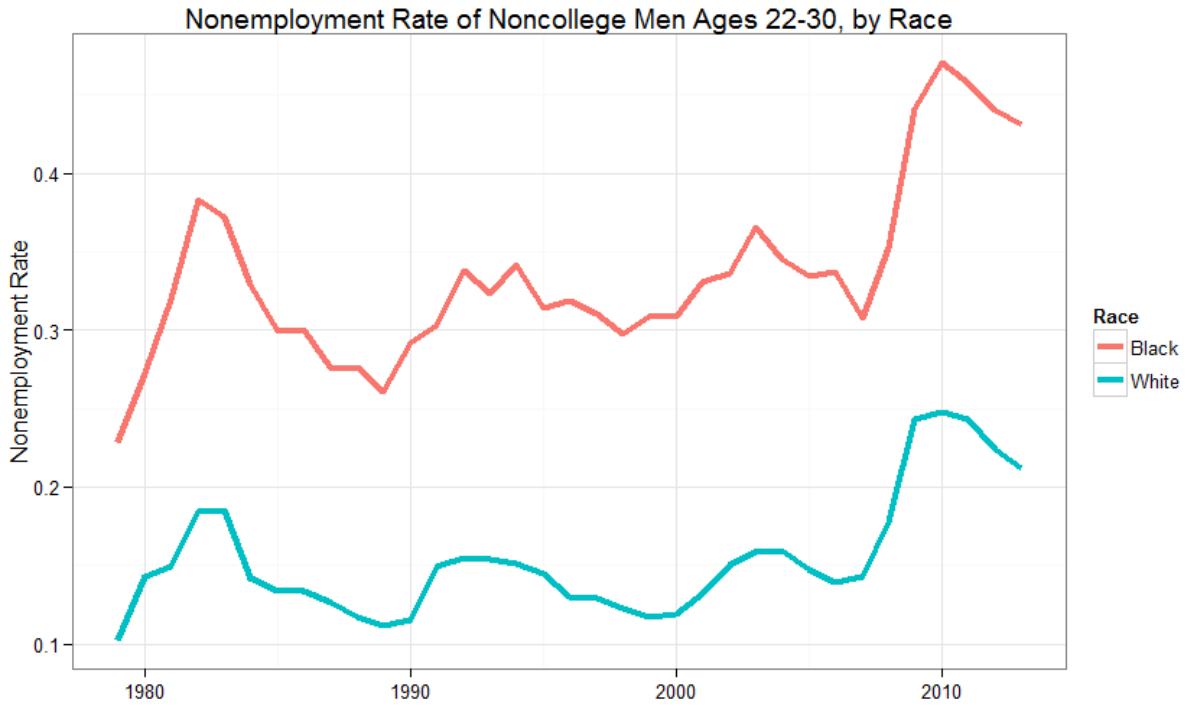


Figure 3.4: This graph compares unadjusted nonemployment rates for whites and blacks, 1979-2013. The rates seem to mostly follow some sort of cycle (possibly the business cycle), especially for whites, until 2009, when both rates suddenly reach record highs. Nonemployment rates decline for both races after 2010, but the rates for each year 2009 or later are higher than any of the rates before 2009. Because both whites and blacks experienced a drastic increase in nonemployment after 2008, and did so fairly equally, the disparity ratios measured in the previous graph are lowest during this period.

3.5 Current Population Survey

The Current Population Survey is probably the most important federal survey of households because it is " is the primary source of labor force statistics for the population of the United States", and it was used to calculate the unemployment rate [167]. It is also used for a number of other calculations, including high school dropout rates, average wages, and voting rates. We examine all of those here, but first we should examine its methodology in more detail. The survey uses 60,000 randomly selected households (which are later weighted[2] to make the statistic more accurate), with each household being in the survey for four consecutive months, out for the next eight, and surveyed again for the following four months. This reduces sample variation and ensures a strong degree of continuity from month-to-month and year-to-year, meaning that any large, short-term changes are more likely to reflect actual changes in the population than random statistical fluctuations. Published data include only populations 16 and older who are neither members of the Armed Forces nor people in institutions (such as prison, nursing homes, and hospitals). Data are collected via a survey of people living in these households, electronically or in person. Suveys include demographic data as well as a variety of other questions, some of which are addressed in questions below. Because of its wide coverage in terms of both population and questions asked, data from the CPS is used by many different organizations for many different purposes. With that in mind, we will examine a few of these issues where the data are particularly affected by excluding prisoners.

3.5.1 High School Dropouts

High school dropout rates are one of the most significant markers of educational progress, mainly because job opportunities are very limited for people without high school degrees (or GEDs) and because not graduating high school usually prevents people from going to college [106]. The reported dropout rate is also quite influential in dictating educational policy, which has been a major part of politics in the U.S. ever since *A Nation at Risk* was published in 1983 and the public was convinced that U.S. school systems were failing. For example, the movie *Waiting For Superman* notes that every president has talked about being the "education president" since that report came out, and many did before as well [77]. As we have seen, the Current Population Survey excludes prisoners because it is household-based. This is easy to adjust for: we simply take the total number of high school dropouts (in prison or free) and divide by the total number of people (in prison or free). Among young black men, we find that the actual dropout rate is about 40% higher than the reported rate, a very substantial difference. Even among young white men (who are imprisoned at a rate about 17% as high as that of young black men), the actual dropout rate is about 11% higher than the measured rate. The fact that imprisonment may directly prevent people from completing high school makes the fact that they hide the number of dropouts even more significant. It is worth noting that the measured high school dropout rates make sense for an employment-focused survey, because it measures the education level of the potential or current labor force (which does not include prisoners). This does not work when it is used to assess how well out education system is doing, though; for that, we need to include prisoners.

[2]Weighting as a techique involves counting some members of the sample multiple times to make it more representative of the population being sampled as a whole. For example, if I were sampling married heterosexuals but my sample contained twice as many females as males, it wouldn't be representative because of a bias toward females. To account for this I may choose to count each male in my sample twice, so that the gender ratio of my sample will equal that of my population, and therefore likely be more representative. The weighting system CPS uses is more complicated because it includes more variables (age, race, etc.), but the process is essentially the same.

Like with the unemployment rate, distortions in the high school dropout rate have grown as the prison system has expanded. This is true both for black and white high school dropout rate and the difference between these rates. Figure 3.5 shows the dropout rates of young whites and blacks, both as reported in the Current Population Survey and as adjusted according to the method described above. The difference between the two rates for whites in minimal (and barely visible before 1998), while the adjusted dropout rate for blacks is significantly higher than the unadjusted rate, and the difference has grown with time.

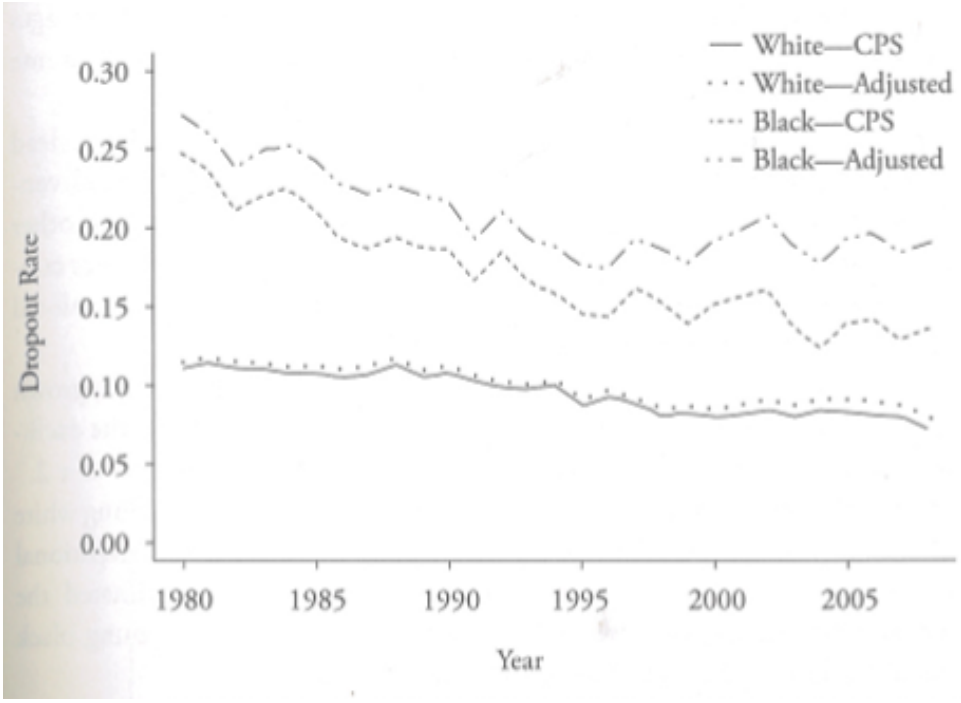


Figure 3.5: *High School Dropout Rate, by Race.* Graph taken from Pettit's *Invisible Men* [126, p. 59]

Figure 3.6 shows the difference between black and white dropout rates, with the measured data represented by the curve below and the adjusted numbers represented by the curve above. Like Figure 3.5, this shows how the difference between adjusted and unadjusted rates has grown over time, so that in 2004 the difference between whites and blacks was about twice as high when prisoners were included.

3.5.2 Wage Inequality

Wage inequality is much more difficult to understand, for a number of reasons. When calculating high school dropout rates it makes sense to include prisoners because people with low education levels are disproportionately represented in prison, and many people are in prison in part because they dropped out of high school (or vice versa). In addition, there is really only one process for including inmates in dropout statistics. Wage inequality is more difficult, and so this procedure needs more justification. If we are interested in the difference in wages between blacks and whites, the easiest way to do this is to compare the average wage for each population. As earlier shown in

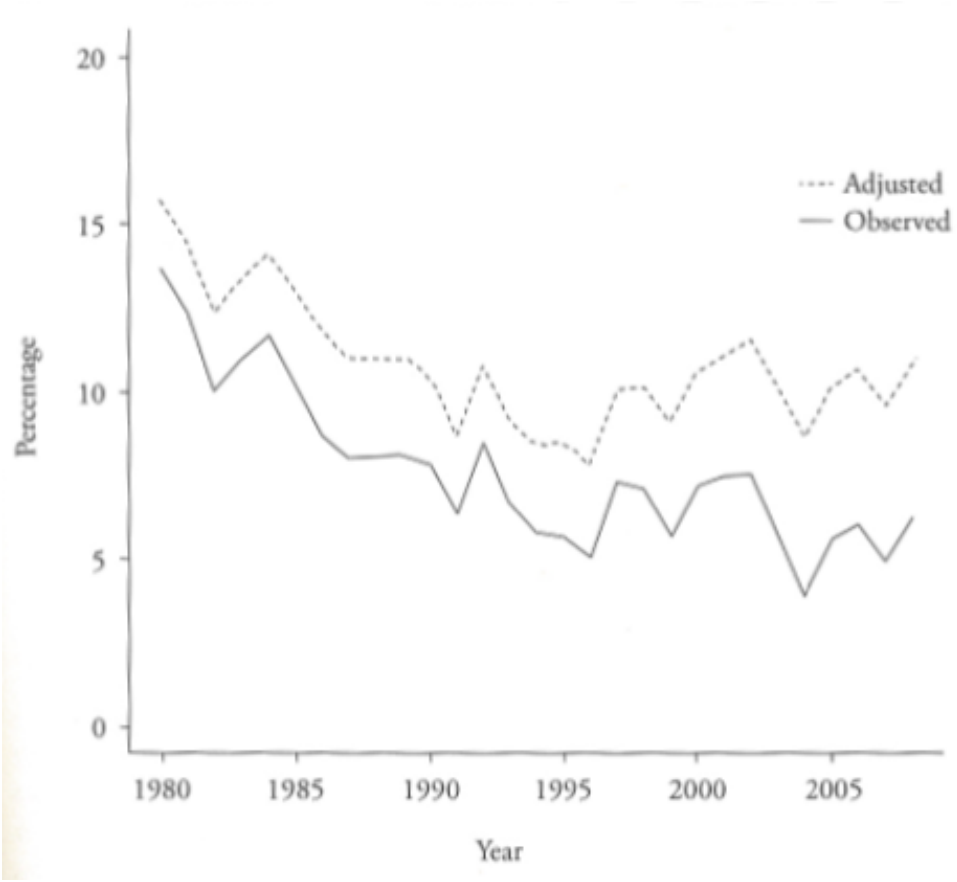


Figure 3.6: *Black-White Difference in High School Dropout Rates.* Also from Pettit [126, p. 61]

Figure 3.4, black unemployment has increased significantly in recent years, raising the possibility that low-paid blacks were pushed out of the labor force (during this same time, the measured racial wage gap declined). If we are interested in looking at discrimination, we do not want blacks being pushed out of the labor market to make the measured wage gap lower, because this implies that the situation is getting better. We could choose to count the unemployed population as earning $0 in wages, which makes sense for certain purposes but does not show anything useful about wage discrimination. For measuring wage discrimination (which here includes lower wages as a result of lower education, lack of experience, etc., provided these are a result of race), it is most useful to look at what the wage difference would be if everyone were employed. The same is true of prisoners; we do not want their being pushed out of the labor market to decrease the measured wage gap. As a result, the statistic we are interested in here is is what the wage difference would be if the entire population were employed. To calculate this, we estimate counterfactual wages for the unemployed and prisoners and recalculate wage statistics with those estimates.

This justification is complicated somewhat by the fact that some prisoners are in fact paid, with wages around 23 cents an hour fairly common. I chose to count them as unemployed when examining unemployment rates previously and make the same choice for looking at wage discrimination, but it is important to give a reason for my choice. In terms of unemployment rates, prisoners with jobs are excluded because the jobs are coerced and the pay is far below the wages of legal employment outside of prison, meaning that on an individual level it is far different from having a legitimate job outside of prison. At the same time, I recognize that in terms of measuring efficiency of economic output it might make more sense to count them as employed since they are producing goods and services. Since I care more about individual benefits of employment than aggregate economic production, and because I think coerced prison laboir has a number of harmful effects that would not be captured in the employment numbers, my default calculation is to exclude them. For looking at wage discrimination, the same justification for calculating counterfactual wages for the unemployed instead of using $0/hour applies. In designing a useful statistic, we do not want prisoners to lower the wage rate by being employed in prison, because we are interested in wage discrimination outside of prison. Wage discrimination should include how different populations would be paid instead of just how they are paid so that it can be measured separately from discrimination in employment and in imprisonment. With these qualitative goals in mind and our methodology well-established and justified, we can move to calculating the statistic that interests us.

Unsurprisingly, the unemployed are often in some important sense less employable than the employed population, meaning that if they were employed their wages would typically be lower. This is true even after adjusting for various demographic characteristics like race, age, education level, and sex. Bruce Western estimates that, within a given demographic category, currently unemployed people would have wages averaging about 80% of employed people's average wage. It is important to do this process by demographic group, because the unemployed are disporportionately poor, uneducated minorities [29]. While it is a good idea in general as a matter of policy, it is especially important in cases like this where we expect the groups to be demographically different. For example, if the average wage of a working 20-year-old (working) poor black male with a high school education is $9.10 an hour, then we assign a wage of $7.28 (80% of $9.10) to all the 20-year-old non-working poor black males with a high school education in our data set. This is repeated for every age/race/sex/education level combination, then the statistic is recalculated. Predictably, the average wage is lower in this calculation, but black wages and white wages decrease by comparable enough amounts that the wage difference stays essentially the same.

Now we can add prisoners to the other unemployed people in our data set. To estimate prisoners' hypothetical wages, we use data from nine correctional surveys reporting pre-incarceration wages from those who were employed immediately before being incarcerated. Like the unemployed, prisoners have qualities that make them less employable even independent of demographic characteristics; on average their wages would be about 42% lower than those of their employed counterparts [176, p. 101]. Using the same process as we used for the unemployed, we assign prisoners hourly wages and recalculate the statistics with our larger data set. This time, the adjustment makes a very large difference, reflecting not only the larger effect of incarceration on hypothetical wages but also the fact that prisoners are more disporportionately black and poor/uneducated than the unemployed. All three lines are on Figure 3.7: the lowest line looks just at the employed population, the middle line at the free population as a whole, and the top line includes the population after adjusting for both nonworkers and prisoners. We can see that the line that includes prisoners is relatively flat after 1985, indicating little change in wage discrimination.

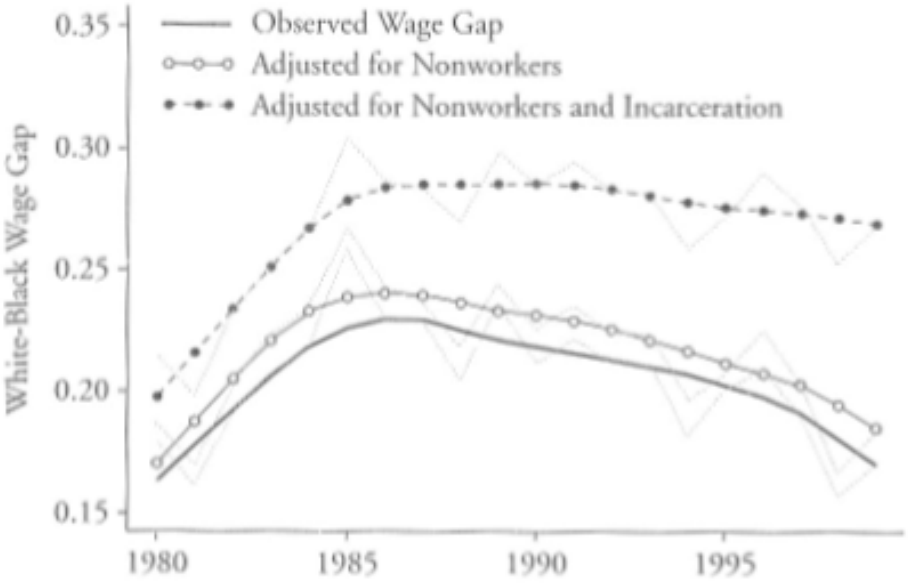


Figure 3.7: *White-Black Hourly Wage Gap.* This graph uses a logarithmic scale, so that the numbers on the left do not directly represent percentages. High numbers still indicate higher levels of inequality. Taken from Western [176, p. 101].

While the original reading of the data was optimistic and showed a declining wage gap, the top line in Figure 3.7 shows that it esentially has not improved since 1985. In fact, it has gotten worse, in that low wages are presumably preferable to unemployment or imprisonment. This is a prime example of how invisible inequality can affect the assessment of public policy, and therefore future public policy. Pettit notes that "[c]laims of improvements in the economic standing of black men, and of young black men in particular, are often supported by analysis of conventional data sources that exclude the incarcerated", citing this as an example [126, p. 67], though Western asserts that the effect is less important for wages. While he notes that "around three-quarters in the apparent gains in relative wages are...due...to esclating rates of joblessness and incarceration", he notes that some scholars are already suspicious of the usefulness of wages to measure inequality given the

visibly higher levels of black unemployment [176, p. 103, 98]. The relation between wages and unemployment means that there is already some awareness around the weaknesses in comparative wage statistics, just as many people are already aware of the limits of the unemployment (versus nonemployment) rate. Imprisonment still augments the effect, however, which is a problem for people who continue to use the wage difference uncritically.

3.5.3 Voting

Like the Current Population Survey, the National Election Study (NES) is household-based, and like unemployment data, it is based on the population of people eligible to vote. In both the NES and the CPS, the 2008 election showed record turnout among blacks, with their voting rate finally reaching parity with whites': 55% of eligible black men voted, compared with 53% of eligible white men [126, p. 78]. This has been attributed to a number of factors, most prominently "Get Out the Vote" Initiatives and Barack Obama being a black man. By using the statistic that looks at the entire population instead of just looking at eligible voters, we can see whether black men did indeed vote in record numbers and were better represented than in past elections or if this is just an effect of parts of the population becoming ineligible to vote. (To be clear, we are actually examining two questions here: civic engagement and political representation. If a lot of people are ineligible to vote but the ineligible population would otherwise vote at very high rates of voting, we can say that civic engagement is high even if political representation is low. Since either measure could be affected by the way statistics are measured and in principle they are at least somewhat independent, a reanalysis can potentially answer both questions.) In this analysis we exclude people who are ineligible to vote because of their age (i.e., minors under age 18) and focus on people ineligible for other legal reasons (i.e. imprisonment and legal disenfranchisement).

First we examine overall representation, looking at what percentage of people are voting (not just what percentage of the eligile population is voting). This process is similar to calculating the non-employment rate: we simply include all people ages 18 and above in the population and recalculate the voting rate. This results in a much lower rate for black men, and in particular black men without a high school education, because these groups are disproportionately disenfranchised. Performing this recalculation, we find that the rate of black voting is not at a record high, and in fact for black men without a high school education is about the same as it was in 1980 20.4% instead of 20.7%). This is especially important because felons and prisoners, when given the opportunity to vote, favor Democrats over Republicans as a group (perhaps because Republicans more zealously promote "tough on crime" attitudes and policies), which means this exclusion disproportionately biases presidential elections in a way that is not fully representative of the population. While multiple elections have been affected, one of the most prominent is from the Bush-Gore presidential race. The election hinged on Florida, which has one of the highest rates of disenfranchised felons in the U.S. Even by very conservative estimates of felon voting tendencies, statisticians have consistently found that allowing felons to vote would have caused Gore to win in Florida, and therefore win the entire election. It is worth noting that during this period, about 1/3 of black men in Florida, or approximately 22,000 people, were permanently disenfranchised because of felony convictions, and that Bush won by only 537 votes [56]. Additionally, 93% of black voters in Florida voted for Gore [122]. These estimates are conservative already, and they do not include prisoners or assume that excluded groups would otherwise vote at rates comparable to the rest of the population. That they still show an almost guaranteed win for Gore indicates how powerful the difference between actual and measured voting rates can be, and therefore how

much statistics that exclude certain populations distort understanding. Christopher Uggen and Jeff Manza note if felons had been able to vote the Democrats may have maintained control of the Senate throughout the 1990s, and this doesn't even include prison or jail imates who were also banned from voting [168, p. 790]. We know that blacks continue to have much less political power than whites as a group, and that this has significant effects in presidential races.

Knowing that blacks are underrepresented politically and that this fact is obscured by the justice system, we can also ask to what extent it is also caused by the justice system. Just as prisoners are on average less employable than non-prisoners, felons and prisoners are less likely to vote than free non-felons, even after controlling for demographic charactersitics [126, p. 74]. We can calculate counterfactual voting rates of the entire population just as we calculated counterfactual wages, and then we can use this as a measure of civic engagement.

Christopher Uggen and Jeff Manza estimate that the voting rate of disenfranchised felons would be approximately 30-40% if they were allowed to vote, depending on the election [168, p. 787]. This is calculated by matching felons with non-felons that have similar demographic characteristics. As predicted, this is lower than the (unadjusted) voting rate of nearly every demographic group analyzed above: the one exception is whites with less than a high school education, who aleady have extremely low voting rates (16.5% in 2008). Using the demographics of disenfranchised felons, the relative voting rates of different demgraphic groups, and the conservative rate of 30% overall, we can estimate voting rates for each demographic subgroup. See Table 3.1.

Table 3.1: *Voting Rates, by Race and Education Level.* This table is recreated from Pettit [126, p. 78]

	Unadjusted	Adjusted	Percentage Difference
1980			
White, college	68.6%	68.5%	0.1%
White, high school	45.8	45.5	0.7
White, less than high school	20.4	20.1	1.5
White, all	55.2	54.9	0.5
Black, college	55.3	54.4	1.7
Black, high school	41.5	39.5	5.1
Black, less than high school	22.8	20.7	10.1
Black, all	42.0	40.0	5.1
2008			
White, college	64.9	64.6	0.5
White, high school	36.4	35.7	2.0
White, less than high school	16.5	14.4	14.6
White, all	53.0	52.1	1.7
Black, college	64.3	63.0	2.1
Black, high school	50.8	46.1	10.2
Black, less than high school	33.5	20.4	64.2
Black, all	55.0	48.6	13.2

Estimating the counterfactual rates of voting among prisoners is more difficult. Vermont and Maine, the two states that allow prisoners to vote at the time of incarceration, do not keep statistics on prisoner voting, but because prisoners are more likely to be uneducated, poor, and disengaged than felons, we have reason to expect that tey vote at lower rates. In the absence of good numbers, we can estimate an upper limit by using the voting rate of felons, and conclude that the actual rate would most likely be lower than that. We can also examine pre-trial jail inmates, who are legally allowed to vote but (with one notable exception[3]) are unable to take advantage of that. We can assume that their rates of voting comparable to or lower than those of felons, although it is unclear whether they would be higher, lower, or comparable. In any case, we will use this for our best estimate of an upper limit to counterfactual voting rates to examine civic engagement. The graph below shows the unadjusted voting rates, as well as rates adjusted for felons and rates adjusted for felons and inmates.

From Table 3.1, we conclude that the voting rate also overstates civic engagement among different populations, although not as much as it overstates political representation. While figure 3.8 shows that the gap between black and white voting rates has narrowed, the rate of voting for whites as a group remains higher than the rate for blacks when adjusted for prisoners and felons. Since the unadjusted numbers show the black rate as higher, this is an important qualitative difference.

3.6 The U.S. Census

The U.S. Census is a unique survey in a number of ways; most important here are its general prominence in research and the special way it treats prisoners. Specifically, prisoners are counted as residents in the town where they are imprisoned. Thus, prisoners contribute to the town's statistics without forming a meaningful part of their communities. At the same time, they are not counted as residing in the communities they lived in immediately before prison (unless they happen to be imprisoned in their own county, which is rare[4]). Because incarceration rates were low until recently, this effect has only become significant in the past few decades, which also is true of virtually all of the statistics mentioned above. While groups like the Prison Policy Initiative are pushing for the census methodology to change, it has not yet and therefore all of the data from recent years suffers from this effect. Distorted data effects measures of growth and other demographic changes, sex ratios, per capita GDP, and other data, both in prison towns and in areas where a lot of people are arrested and sent away to prison. This paper examines some of these effects and, where possible, how to adjust for them. In many cases, however, the data simply do not exist (or are extremely difficult to find), so that even a knowledgeable and committed researcher might not be able to get the correct information. Thus, while understanding and examining the issue here will improve our understanding of census data, ultimately we will not be able to get all the data we want without a change in the census methodology.

This section uses the term "prison county" to refer to a county where a large portion of the

[3]Jails in Washington D.C. set up polling stations with inmates' absentee ballots, so that the experience is fairly similar to voting at most voting precincts. Rates of voting are still quite low in these jails; in 2012 88 out of 1,700 inmates voted in the election [74].

[4]As noted above, prisons tend to be built in isolated, rural areas. Since prisoners disproportionately come from populated areas (the crime rates in cities are higher, and populated areas have more people), and also because prisoners are often frequently moved between different state and federal prisons, most prisoners are not incarcerated where they were arrested [131].

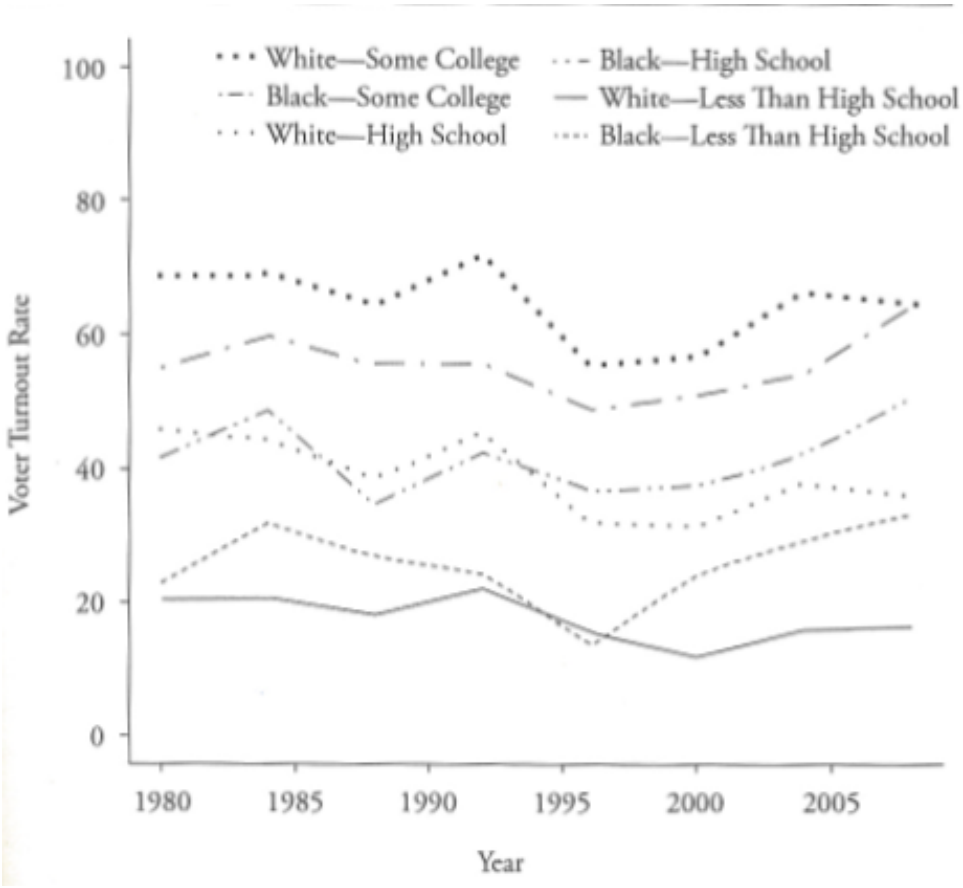


Figure 3.8: *AdjustedVoter Turnout Rates Over Time, by Race and Education Level.* This graph shows that the voting rates of blacks relative to white have been increasing in every category. Interestingly, this graph also shows how important education level is in every year; each voting rate for high school dropout is lower than all of the voting rates for people who completed high school, which is lower than any of the rates of people who completed college. Graph taken from Pettit [126, p. 80].

measured population is made up of prisoners, and the term "arrest county" to refer to a county where a large portion of the (legal resident) population is incarcerated in another area. To illustrate some of these distrotions, as well as their importance, I use Brown County, Illinois (a rural county) as an example of a prison county and Cook County, Illinois (an urban county that includes Chicago) as an example of an arrest county throughout this section. Bear in mind that the prisoners in prison counties like Brown come from arrest counties like Cook; thus, the changes in these two types of counties are interdependent.

Unless otherwise noted, the data from this section come from the Prison Policy Initiative's report *Too Big To Ignore* [130].

3.6.1 Overall Population Trends

The most basic demographic information about a county is its population. Prison counties will always have inflated measured populations, while arrest counties will always have lowered measured populations. In Brown County, the population is highly overestimated, since 28% of the measured population consists of inmates. The population of Cook County, on the other hand, is understated, although it is hard to say exactly how much. The population distortions here and in other counties are immediately significant because (among other reasons) many allocations are made on the basis of population by the federal government. Perhaps most significant among these involves the congressional districting system and how votes are counted. Voters in prison counties (or in districts containing prison counties) will have more influence than normal under the districting system, and voters in arrest counties will have less. As we saw in the previous section, prisoners are not represented politically on an individual level (except in Vermont and Maine); we now see that communities damaged by the prison boom are additionally underrepresented at a group level. Prison counties, which are likely contain many prison towns whose economies are apparently dependent on the prison industry, are also overrepresented. Together, these two trends lead to public policy that at least encourages maintaining the prison population at its current level, if not increasing it.

Population is a static statistic, meaning it focuses on a point in time. Changes in population level, or growth and decline, are the corresponding dynamic statistics and also suffer from population distortions. Some 56 different prison counties[5] appear to be growing and healthy, despite the fact that their (unincarcerated) populations are actually declining (see Figure 3.9). Brown County is not one of these, but 7 other counties in Illinois do exhibit this trend. Their prison population is expanding sufficiently quickly that the census reports them as growing, even as their legal resident population is declining. On the other hand, if Brown County were to close all of its prisons its reported population would decline by more than 25%, which would make it look extremely unhealthy. Similar effects can be seen whenever a prison closes[6]: a prison county that has a prison close will appear to suffer from a sudden, precipitous decline in population even if the actual population is stable (or growing slightly). These effects are easy to adjust for; simply subtract the number of prisoners from the total population and use that as the new measure of population. Since having prisons can be harmful to communities in a number of ways in terms

[5]Of these 56, 53 are classified as nonmetro counties while only 3 are classified as metro counties. Since prisons tend to be rural, prison counties showing the effects discussed in this section will be largely rural as well, and rarely metro.

[6]For example, the Census Bureau's 2002 estimate of the population of Mahoning County in Ohio reported a decrease of 4,250 people; roughly half of this decrease is due to the closing of a large private prison.

of disrupting the local economy, having a prison close down might actually be very positive for a community's health (aside from causing a loss of jobs in the short term), but the census will make it look as though the community is in great danger by reporting a drastic decline in population. These effects can be powerful at the level of entire populations; when looking at a smaller proportion of the population (e.g. a racial group), the effect can be much more extreme.

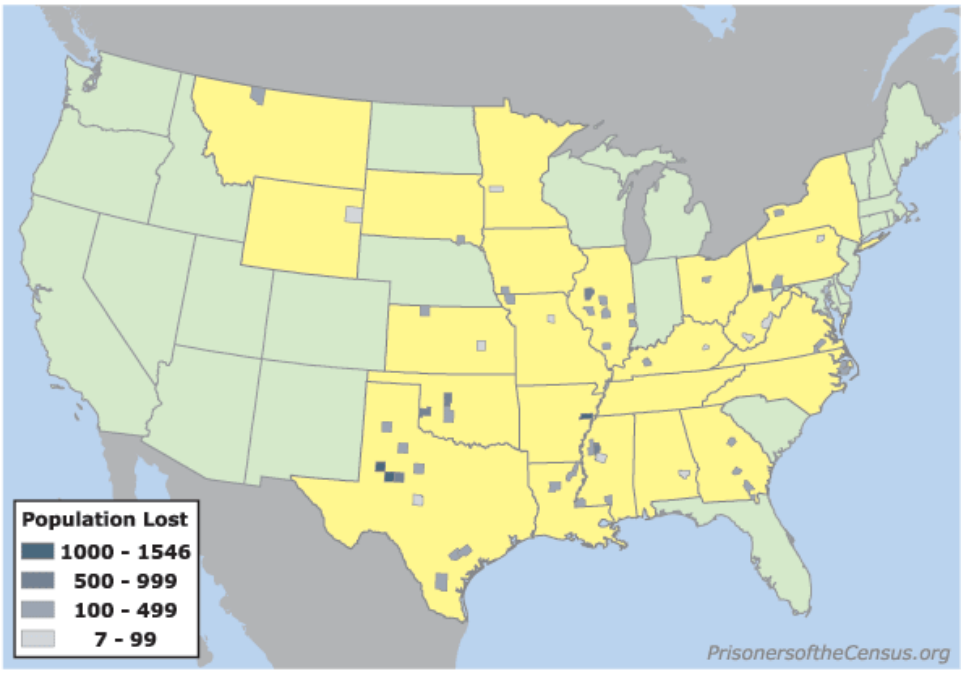


Figure 3.9: *Counties With Positive Growth Rates in the Census, but Shrinking Free Populations.* This shows which counties experienced growth only because of an increase in prisoners, and how much their free population decreased during that period. Taken from *Too Big to Ignore* [130].

3.6.2 Race, Ethnicity, and Gender

Given the racial, ethnic, and gender differences between prisoners and the general population, it is not surprising that census data are more distorted for some of these groups than for populations as a whole. In 2004 there were 173 counties where the majority of black residents were prisoners; now[7] there are 206, 20 of which are in Illinois. In five of these counties the entire black population is incarcerated (although none of them contain more than 15 black prisoners, or more than 100 prisoners total). Brown County, on the other hand, contains 1,260 incarcerated blacks, and only 5 who are not incarcerated. Thus, 99.60% of its black population is incarcerated, as reported in the Census. In Cook County, on the other hand, less than 1% (0.65%) of the black population is reported as incarcerated; prisoners therefore contribute very little to Cook's measured population numbers, despite the fact that most prisoners in Illinois come from Cook.

While not as drastically as blacks are, Hispanics and Latinos are also disproportionately incarcerated and also suffer distortion effects. Brown County has a high proportion of its Latino

[7] As of March 26, 2014.

population incarcerated at 86.45%, with 236 Latino prisoners and 37 Latino non-prisoners. More significantly, there has been a movement in recent years of young adult Hispanics from urban into more rural areas. Because the counties that have a large proportion of young adult Hispanics are mostly the same counties that incarcerate large numbers of Latino prisoners, it is unclear whether this movement is in any way voluntary or whether it is entirely due to the census measuring procedures (and most prisons primarily housing formerly urban residents in rural areas). Without looking at and adjusting data from every individual prison county, we do not know whether the trend is meaningful outside of imprisonment and have no way of finding out.

Finally, we examine gender ratios. Because 92% of prisoners in the U.S. are male, this effect is especially powerful. In prison counties (and counties with military bases) males are vastly overrepresented, whereas in arrest counties they are severely undercounted. In prison counties it is usually possible to determine what the non-prisoner sex ratio is, by subtracting the number of male prisoners from the male population (and females from the female population) and then recalculating the ratio. In arrest counties, however, it is nearly impossible, because the census does not keep data on the number of people from that county sent out of state. (Even when the information is available at the county level, it will almost certainly not be available at other levels, like the neighborhood level.) Even more difficult than calculating gender ratios are difficulties in calculating the proportion of the population that is married. Many government programs are dependent on data about the size of the married population (or a married subpopulation), and statistical measures like the ratio of unmarried males to unmarried females are doubly hard to calculate. The measured number of unmarried men per 100 unmarried women in Brown County is at least 165.5, while the same ratio in Cook County is 85.5 or lower[8]; it is entirely unclear how close these two numbers would be if prisoners were counted in their original communities. Figure 11 shows the rate of unmarried men per 100 unmarried women across the U.S.

Basic demographic data about gender is used in nearly all types of social science research, which is what makes the census such a valuable survey. That it misrepresents the truth, often grossly, and in many cases does not enable researchers to find the true data they are looking for, is therefore significant for all fields of social science research and public policy.

3.6.3 Income Statistics

We have already seen how prisoners affect income statistics in terms of wage equality in the CPS. Using the census data they affect income statistics in another way, by distorting per capita income. In prison counties, the measured income per capita is artificially low because prisoners earn no (or negligible) income but contribute to the population measure of a county. To the extent that people making use of per capita statistics are interested in measuring the well-being of everyone who resides in that county this actually makes sense, but if per capita income in meant to be a measure of the economy in the free community (which it usually is), this causes prison counties to look much worse off than they are. This error is easy to fix, fortunately: simply multiply the measured per capita income by the number of residents *including* prisoners, and then divide by the number *excluding* prisoners. Like all of the statistical measures talked about in this paper this is still a problem because the corrected statistics are rarely used, but an interested researcher is able to easily fix this.

[8]The map this is based on classifies Brown as 165.5-362.1, and Cook as 53.8-85.5. So the number in Brown could be much higher, and the number in Cook much lower, than those given in my original sentence.

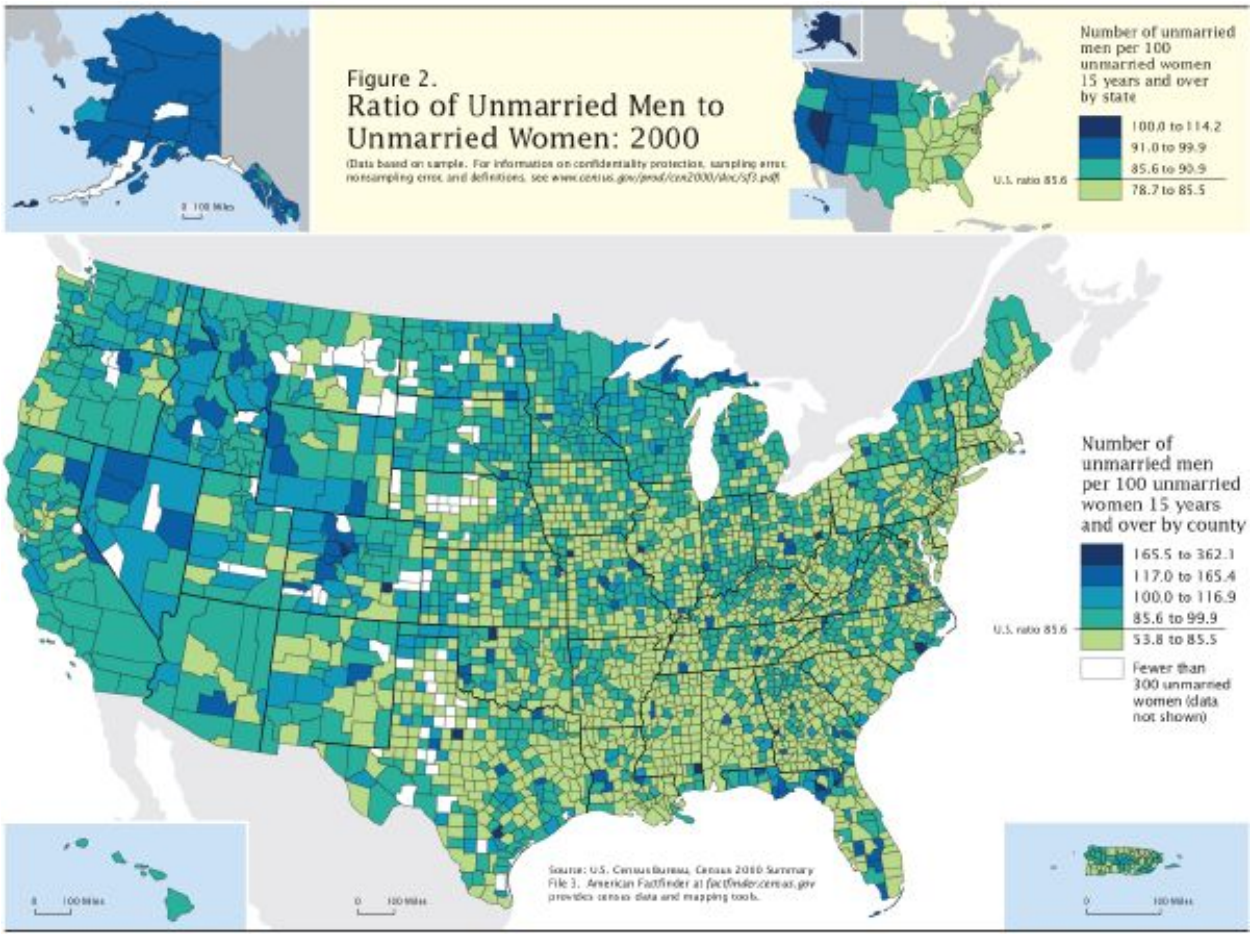


Figure 3.10: *Ratio of Unmarried Men to Unmarried Women in 2000.* The darkest county in Illinois is Brown County, while the light county in the Northeast corner of Illinois is Cook County. Taken from *Too Big to Ignore* [130].

For arrest counties, the problem is more nuanced. In a sense the measure is appropriate as it is; it measures the average free income of residents there and in that sense is probably the best way to describe the economic activity of a community. Since this is probably the most common use of per capita statistics, this measure seems well-chosen. At the same time, it is worth remembering that income statistics are also used as a measure of health of the community. An arrest county, in that it has a lot of people sent away to prison (and in theory a lot of crime), is not nearly as healthy as looking only at the free population would imply. More specifically than health, per capita income statistics can be seen as the expected future earnings of a random person in that community. Since we are looking at arrest counties, by definition many of the people are going to be sent to prison (if present trends continue), meaning the average expected life earnings are actually significantly smaller than the statistics imply directly. Like most statistics involving arrest counties, this is difficult to fix for a given arrest county, and virtually impossible for a small arrest neighborhood.

3.7 Health

Aside from the Current Population Survey and the Census, most major federal surveys are about health [126, p. 29]. Almost all of these health surveys are also household-based, and therefore exclude prisoners and some other populations [126, p. 29]. This sections looks at two of these surveys, the National Epidemiologic Survey on Alcohol and Related Conditions and the National Survey on Drug Use and Health, and how excluding prisoners in these surveys affects estimates of alcohol and substance disorders in the DSM-IV.[9] While prisoners are known to have high rates of a number of phyiscal and mental health conditions (and therefore health statistics would be relevant in a number of areas), alcohol and substance disorders and particularly significant in that many people are in prison specifically because of their alcohol or substance use [125, p. 46-50]. Prisoners house some 877,000 inmates with alcohol use disorder and 1,043,000 with substance abuse disorders, none of which are counted [36]. Perhaps in part because alcohol use is a status crime (meaning that use is legal after a certain age) while illicit substance use is not, substance abuse estimates suffer much larger distortions.

Including prisoners raises the measured prevalence of alcohol use disorder by about 5%, with a 4.2% increase for alcohol abuse disorder and a 5.9% increase for alcohol dependence disorder [36]. All of these changes are notable and relevant to people who utilize drug use data, although because the numbers are relatively small they generally do not matter very much. For substances, the reported rate of illicit drug use disorders increases by 25% when prisoners are included, with a 12% increase in illicit drug abuse and a 53.8% increase in illicit drug dependence [36]. These estimates are large enough that the difference is practically important, and knowing that drug use is concentrated in the prison and jail population can allow more effective treatment of drug problems. As with all of these statistics, effects that are important generally become more so for certain groups: cocaine and heroin are uncommon outside of prison (at least compared to alcohol) but very common in prison, meaning the difference likely to be even larger for them than for illicit drugs generally [36]. The rates for poor minority groups (who are over-represented in prison) are likely to be particularly affected because there are proportionally many more excluded prisoners.

As a note about health statistics in prison in general, health is a particularly interesting issue

[9]The DSM-V came out this past May, so the numbers below will be slightly different, but since the methodology is still household-based the basic effect should be the same. The numbers are from 2010 and would be slightly different in any case.

in prisons for a number of reasons. Prisoners have higher rates of a variety of physical and mental conditions, both because they come in like that and because prisons often directly cause them. Moreover, prisoners represent a significant proportion of people with mental illnesses (and certain physical illnesses[10]); a recent report notes that the number of mentally ill people in prisons and jails is more than ten times as high as the number in mental hospitals [138]. Because prisoners are a captive population directly under state control, they would be in the best position to provide treatment if states were more serious about treating prisoners. As prisoners are later released into their communities and often have few resources for dealing with health conditions (and in the case of physical illness, often infect other people), focusing more on prisoners would help reduce rates of a lot of health conditions in the general population. Even if policymakers do not care about the health of prisoners directly, they would benefit from seeing health conditions in prisoners and realizing that when those prisoners are released these will become health conditions in the general populace. Insofar as statistics include prisoners, distributions of health conditions will highlight the prominence of prisons and likely focus epidemiological and other health measures on prisoners more.

3.8 Conclusion

Excluding prisoners from national surveys distorts the understanding of the public and policymakers for a number of social issues. In the Current Population Survey or anything else household-based that measures statistics only within the eligible population, the exclusion of prisoners will affect the data, sometimes quite significantly. Effects significant at the national level are even more so at the community level for states, counties, cities, and neighborhoods where large numbers of people are incarcerated. We have seen significant effects on measures of unemployment, the dropout rate, racial wage inequality, and voting rates, and that in all of these cases the measured statistic gave a more optimistic view of the social problem than the adjusted statistic. In the census data we see that the practice of counting prisoners as "residents" of the communities they are imprisoned in affects estimates of population, race and gender ratios, and income statistics, although in this case the change could be favorable or unfavorable. In any case, we have seen that the distortions on these two surveys alone are enormously significant, not only for excluding prisoners but also because they often fail to account for felons or ex-felons and people who are homeless or have unstable housing. That many other surveys also show these limitations shows the importance of accounting for "invisible inequalities". While few surveys individually rival the CPS and census in use, the overall number of other surveys that produce similar effects is enormous, and the combined impact of the distortions resulting from these other surveys is infinitely greater than the impact from the CPS and census alone. We therefore see how essential an understadning of prisons is to the understanding of other social problems, or even just social statistics (like marriage rates). Perhaps one of the greatest impacts prisons have is altering people's perceptions of other social

[10] Joan Petersilia notes in particular that prisoners have high rates of HIV/AIDS, hepatitis B and C, and tuberculosis, with roughly 38% of U.S. tuberculosis cases coming from prisoners [125, p. 50]. In this particular case it seems that prisoners are included in at least some of the statistical measurements, but three points are still worth noting. First, it is likely that statistical measures on some other physical ailments (as well as some other measures of tuberculosis) do not include prisoners, and these measured differences have shown that the difference can be large enough to matter. Second, because rates of homelessness are much higher among former prisoners, it is likely that many former prisoners are not being counted. Third, focusing more directly on the number of prisoners with these diseases allows for more efficient public health interventions.

problems, and therefore their individual and public policy responses. It is therefore essential that these statistics are corrected, so that any number of social problems can be addressed.

At the same time, we need to remember that the methodological choices for calculating these statistics is intentional, and for their primary intended purposes is probably the most useful. Insofar as unemployment rate is used primarily to describe bargaining power of workers or the high school dropout rate is used to predict the productivity of future workers, the system as set up is appropriate. As soon as these measurements are used for other purposes, though, there is a risk that these choices will prove unhelpful, and for many purposes they clearly fail to be optimal. When working with CPS data or looking at prison communities, it is usually fairly straghtforward to adjust the data, as prisoners can be added or excluded as appropriate. For arrest communities, however, data are frequently extremely difficult (or even impossible) to adjust as may be appropriate. I believe that a change should occur in academics to emphasize using the adjusted statistics (at least providing both and making it clear to people what each is used for), but regardless of how that works the census data needs to be collected in a new way. The old method, so useful for many types of information, fails when dealing with issues that that rely on the other type of data. I think the census data should by default measure prisoners as members of the the communities they resided in before being arrested, but either way it should certainly provide people with the data needed for researchers to analyze the data either way. If the census makes it possible for people to adjust their data and a culture of doing so forms among academics, people's understanding of social problems will be much improved. This will lead to better intervention and ultimately a better world.

Methodological Appendix – Calculating Nonemployment Statistics

The first step for calculating adjusted nonemployment disparity rate was calculating nonemployment rates for blacks and whites using CPS data from each year. Because CPS data was available from 1979 through 2013, I calculated this rate for each of them. Following Western's methodology, these statistics were calculated only for noncollege males ages 22-30, for single-race whites and single-race blacks, with a comparison between the nonemployed part of the population and the population as a whole. An example year, 1979, is below with comments:

```
#Data -- 1979 (example year)
library(foreign) #This allows R to use the read.dta command for the dta CPS files
data1979 <- read.dta("morg79.dta")
data1979a <- data1979 #Creating a duplicate is faster than loading the
#entire dataset into R twice

#This limits the sample size to males; I exclude NAs for all these variables
#because their meaning is unclear
data1979a <- data1979a[data1979a$sex == "Male", ]
data1979a <- data1979a[is.na(data1979a$sex) == FALSE, ]

#Limits the sample to people ages 22-30
data1979a <- data1979a[data1979a$age < 31, ]
data1979a <- data1979a[data1979a$age > 21, ]
data1979a <- data1979a[is.na(data1979a$age) == FALSE, ]

#Here I consider only blacks; the code is repeated for whites using data1979a
black1979 <- data1979a[data1979a$race == "Black", ]
black1979 <- black1979[is.na(black1979$race) == FALSE, ]

#This limits the sample to people who have not completed any year of college
#and aren't currently in school.  This isn't perfect, because they could be
#currently in high school.
black1979 <- black1979[black1979$gradeat < 13, ]
black1979 <- black1979[is.na(black1979$gradeat) == FALSE, ]
black1979 <- black1979[is.na(black1979$doinglw) != "At School", ]

#The variable here measures labor activity; this gives the number in the
#total population so I can calculate the ratio later.
black1979 <- black1979[is.na(black1979$ftpt79) == FALSE, ]
nrow(black1979)

#This eliminates people who are employed from the sample, leaving only
#the nonemployed
black1979a <- black1979[black1979$ftpt79 != "Employed full-time", ]
black1979a <- black1979a[black1979a$ftpt79 != "Part-time For Economic Reasons", ]
black1979a <- black1979a[black1979a$ftpt79 != "Employed PT", ]
black1979a <- black1979a[is.na(black1979a$ftpt79) == FALSE, ]
nrow(black1979a)
```

```
#This gives the nonemployment rate for blacks (number of nonemployed
#divided by total population)
nrow(black1979a)/nrow(black1979)

#This is then repeated with whites, starting with
data1979a <- data1979
#and being different only in this line and subsequent use of the variables
#white1979 and white1979a
white1979 <- data1979a[data1979a$race == "White", ]
```

This gives me the rate of unemployment, by race, for the year 1979; it is repeated with each other year. The variables changed a bit from year to year; changes included the reclassification of some variables as numbers (e.g., sex=1 for males, race=1 for single-race whites), changes in years (gradeat became grade92, starting in 1992), and the change from ftpt79 (measuring whether work was full-time, part-time, or neither) to lfsr94 (directly measuring employment, unemployment, and not being in the labor force). The full R code is included in the electronic file for anyone interested in reproducing my work exactly.

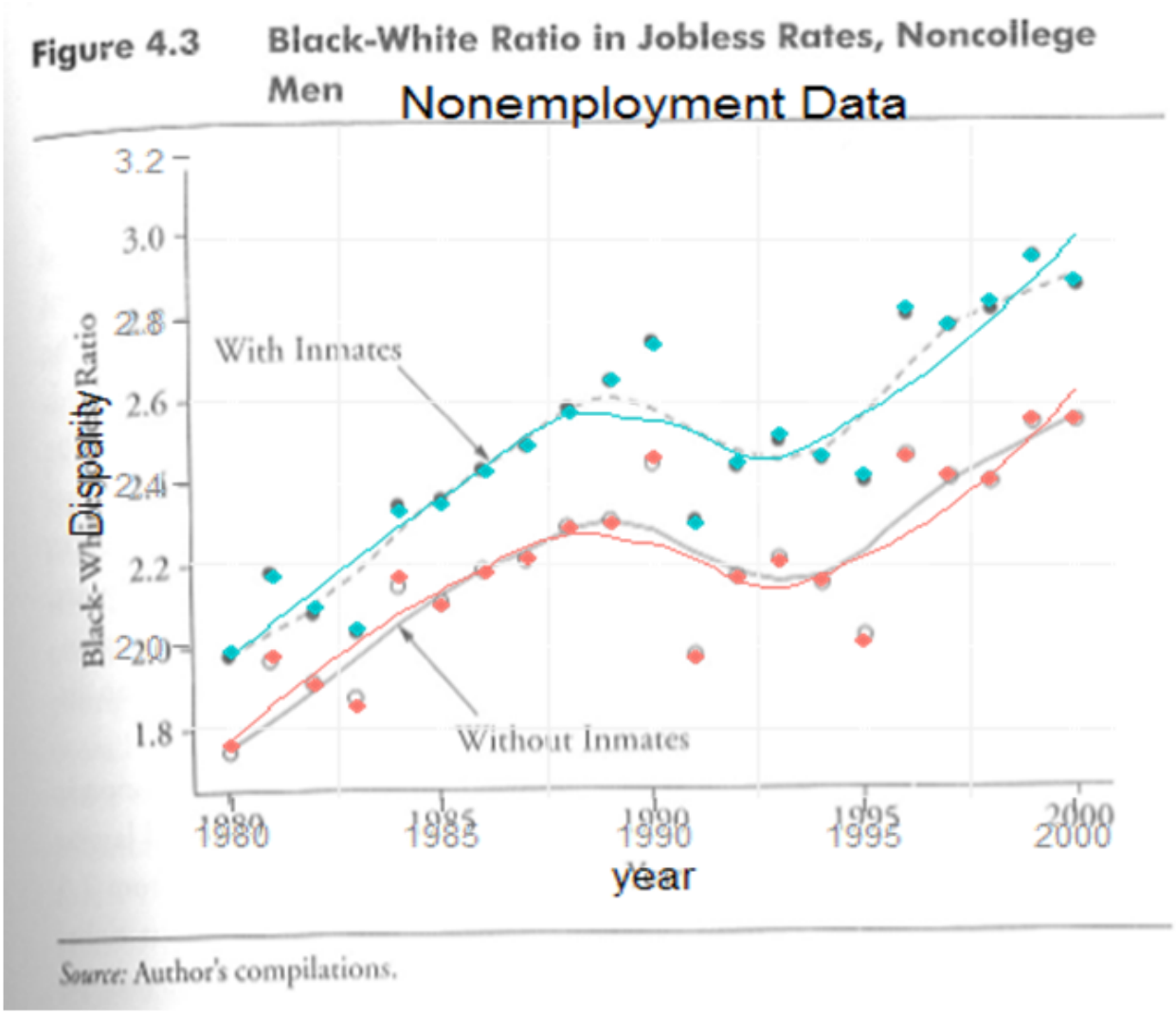
Most of the rest of the work was done in excel (excel file also attached). The unadjusted disparity ratio was calculated by dividing the rate for blacks by the rate for whites. Calculation of the adjusted ratio was a bit more complicated. Data from reports named "Prisoners in 2000", etc. from the Bureau of Justice Statistics give number of prisoners per 100,000 by race, sex, and age combinations, as well as total number of prisoners in those categories. Their age categories didn't line up perfectly, however (they were 20-24, 25-29, 30-34, and other ones irrelevant to this analysis). I used the number for prisoners ages 25-29 directly, then used 4/5 of the number ages 20-24 to estimate the number for ages 22, 23, 24, and 30 (this is a conservative estimate for age 30 since the number for 30-34 was consistently than that of 20-24, but it reports the 22-24 data more accurately and I prefer to be slightly conservative in my statistical estimates). With the total number and the rates in each category, I calculated the total population of black/white males ages 22-30 and divided the number of prisoners by that number to get the overall rate [an example column in the excel document looks like this (E2+F2)/(E2/I2+F2/J2)]. (This works better than just averaging the rates directly, because the size of the total U.S. population ages 20-24 might be significantly different from the size of the population ages 25-29.) I then calculated the adjusted rate of nonemployment for each race [(1-M2/100000)*B2+M2/100000 ; I multiply the unadjusted nonemployment rate by (1-M2/100000) rather than simply add the rates because the CPS doesn't count prisoners in the total population or in the nonemployed population], and finally the adjusted ratio.

Unlike Western, I did not attempt to adjust for jail inmates or reclassify military personnel as employed. This was for two main reasons: first, I couldn't find reliable data and jail inmates (i.e., it was less specific than prison data, and it didn't go beyond 2002[11]), and also because the number of black male jail inmates is roughly equal to the number of black military men (and both populations are skewed towards younger men) so my prediction is that they would roughly cancel each other out.

[11]This is true of the BJS data but not the Census data. However, the census data is even less specific, with broad age categories like 18-55, and it involved more estimation than I thought acceptable for the level of certainty I prefer in my claims.

Code for Creating Graphs in R

```
library(ggplot2)
#My first graph is a recreation of Bruce Western's graph in Punishment and
#Inequality in America.  Before creating the graph featured in my paper, I
#estimated his number and then created this picture, which overlays his
#original graph with a graph I created by using estimates in the data.  Here's
#the comparison picture; as you can see both the points and the loess lines
#are close matches, and the scales are aligned:
```




```
#Below is the code used to make this graph, which is also the one
#used in my paper.  This csv includes both the data from my analysis
#of CPS data and the numbers from Western, so this first command
#restricts it to data from Western for this graph:
data <- read.csv("Western.csv")
```

```
data2 <- data[data$Author == "Western", ]

ggplot(data2, aes(x=Year,y=Disparity,color=Data)) +
geom_smooth(fill="white") +
geom_point(size=3) +
ggtitle("Black-White Disparity Ratio in Joblessness, Noncollege Men Ages 22-30") +
scale_x_continuous(limits = c(1980,2000)) +
scale_y_continuous(limits = c(1.7,3.1)) +
xlab("") + ylab("Black-White Disparity Ratio") +
theme_bw()

#For my next graph, I compare my CPS-generated data with
#Western's.  The syntax is differet because I do this as five
#separate plots, layered on top of each other:
data <- read.csv("Western3.csv")
ggplot(data, aes(x=Year)) +
  geom_line(aes(y=data$Western),color="Green1",size=1.5) +
  geom_line(aes(y=data$Kalla),color="Red1",size=1.5) +
  geom_smooth(aes(y=data$Western),color="Green4",size=1) +
  geom_smooth(aes(y=data$Kalla),color="Red4",size=1) +
  geom_point(aes(y=data$Adjusted),color=data$Author,size=3) +
ggtitle("Disparity Ratios in Nonemployment, by Western and Kalla") +
scale_x_continuous(limits = c(1980,2000)) +
scale_y_continuous(limits = c(1.7,3.2)) +
xlab("") + ylab("Black-White Disparity Ratio") +
theme_bw()


#This graph includes the full range of my data.  The biggest change
#from before is that the limits of the graph have changed.  It's
#also uploaded from a different csv file; while my original csv
#file had all the data I use, I found it more convenient to create
#additional files, with each one being optimally formatted for
#whichever graph I was making.  The other main change here is
#the removal of loess lines.  Code below:

data <- read.csv("Fulldata.csv")

ggplot(data, aes(x=Year,y=Ratio,color=Type)) +
#geom_smooth(fill="white") +
geom_line(size=1.5) +
#scale_shape_manual(values=c(10,19)) +
ggtitle("Disparity Ratios in Nonemployment, 1979-2013") +
scale_x_continuous(limits = c(1979,2013)) +
scale_y_continuous(limits = c(1.7,3)) +
xlab("") + ylab("Black-White Disparity Ratio") +
theme_bw()
```

```
#For my final graph, I look at the nonemployment rates of
#whites and blacks and compare them.  Like the previous
#graph, this lacks loess lines (largely because the change in
#employment around 2008 was not smooth, and loess lines
#push the change back to 2000; I also use a different csv
#file with a different set of data:


data <- read.csv("RacialNonemployment.csv")

ggplot(data, aes(x=Year,y=Rate,color=Race)) +
#geom_smooth(fill="white") +
geom_line(size=1.5) +
#scale_shape_manual(values=c(10,19)) +
ggtitle("Nonemployment Rate of Noncollege Men Ages 22-30, by Race") +
#scale_x_continuous(limits = c(1980,2000)) +
#scale_y_continuous(limits = c(1.7,3.1)) +
xlab("") + ylab("Nonemployment Rate") +
theme_bw()
```

Chapter 4

No Punishment Intended: A Proposal for Restorative Classroom Management

Abstract

In this paper, I propose a three-part program for classroom discipline to help schools deal with disciplinary problems in the classroom. Based on a review of academic literature, my thesis is that the way a teacher deals with classroom discipline should be immediate, certain, and non-punitive, and that it should keep students integrated into the classroom environment, empower students to become more responsible, and make teaching easier for teachers. It includes a method for handling disruption in class, a mini-curriculum to teach students about conflict resolution, and some professional development for teachers. After describing this proposal I discuss some concerns relating to implementation of the policy. I conclude by looking at how the effectiveness of the policy will be measured.

4.1 Background and Literature Review

Introduction

Student discipline is a major concern for all teachers, and it is one of the most important parts of running a classroom. The National Education Association notes how essential it is at all levels [182], and education writer Jay Mathews notes that for beginning teachers struggling with low-performing students, "the issue is the teacher's failure to maintain discipline." [109] As teachers struggle to maintain effective classroom discipline, schools across the U.S. have increasingly adopted zero tolerance and other harsh policies for students, with the result that suspensions, expulsions, and even imprisonment are becoming increasingly common for youth in schools.[1] It is therefore to the benefit

[1] A *Times* article from 2012 notes that "[s]uspension rates have more than doubled over the past three decades across all grade levels. At the same time, racial gaps have widened: black students are 3 1/2 times as likely to be suspended or expelled as their white peers, according to Department of Education data" [32].

of both teachers and students if classrooms can be managed more effectively so that teachers do not feel the need to resort to harsh punitive measures to maintain control. I will briefly review the literature on the prevalence of these policies before addressing the causes of these shifts and how they relate to other trends encouraging punishment. Afterwards, I discuss the effects of these policies, in terms of psychological theories of punishment and sociological theories of how poverty and criminality are perpetuated. I conclude by discussing how this informs my policy proposal. Specifically, I conclude that the way a teacher deals with classroom discipline should be immediate, consistent, and non-punitive, and that it should keep students integrated into the classroom environment, empower students to become more responsible, and make teaching easier for teachers.

4.1.1 Causes of Zero-Tolerance and Policing

From reviewing the literature I have identified four primary causes of the growth of zero tolerance policies and policing in schools: broken-windows theory, the crime wave of the 1970s and 80s (along with the media image of juvenile superpredators and the War on Drugs), the rise in surveillance as central in society, and the additional pressures on teachers for efficient classroom management resulting from No Child Left Behind and the school privatization movement. In addition, persistent inequalities in schools and large class sizes has exacerbated these trends, which have been most extreme in low-income, minority schools. Each of these four factors will be discussed in turn.

4.1.1.1 Broken-Windows Theory

In 1969, Phillip Zimbardo conducted in experiment in which a car was left abandoned with the hood up in the Bronx, NY and in Palo Alto, CA. The car in the Bronx was almost immediately broken into and stolen from, with everything of value stripped within 24 hours, while the car in Palo Alto was left unharmed for a week. Zimbardo smashed part of the car with a sledgehammer and left it in Palo Alto again; it was quickly destroyed and looted [111, p. 64]. Zimbardo concluded from this experiment that social context is a major determinant for behavior; the run-down environment of the Bronx and the pre-existing smashing of the Palo Alto car were cues to people that they could destroy and steal from the car, while the un-smashed car in Palo Alto had provided no such cues and as a result was unharmed after a week. This, along with a case study of policing in New York, became the basis for broken-windows theory, proposed in 1982 by James Q. Wilson and George L. Kelling. There had been a series of violent crimes in the NY subways, and police were unable to prevent them. They tried a counter-intuitive strategy of focusing their efforts of reducing *low-level* deviance (such as graffiti, not paying fees, etc.), with the idea that crime resulted form the appearance of neglect, and that by cleaning the graffiti each night they could show that the subways were being cared for and thereby prevent serious violent crimes. The crime rate subsequently fell, drastically. On this basis, Wilson and Kelling proposed their broken-windows theory, which states that signs of neglect (like broken windows) indicated a weak response to crime and therefore induced more criminal activity.[95] They essentially agreed with Zimbardo, but more directly applied the importance of social context to crime and the idea of public-order policing.

Because broken-windows theory requires consistently punishing even low-level deviance, it later inspired the growth of no, punitive 'zero tolerance' policies in schools [27, p. 46].[2] Zero toler-

[2]There has been some scholarly disagreement on the tie between zero tolerance and broken-windows theory, in part because of the philosophical differences between them. As noted by the American Psychological Association's Zero Tolerance Task Force Report, "It is important to note that 'broken window theory' as promulgated by Kelling

ance policies are policies that "assign...explicit, predetermined punishments to specific violations of school rules, regardless of the situation or context of the behavior," with the punishment often being quite harsh [20, p. 1]. Proponents of zero tolerance policies note how they send a strong signal to any potential rulebreakers, while also removing people who have broken rules in the past from the school population so they do not continue to model rulebreaking with impunity. By serving the double purpose of showing little tolerance of misbehavior *and* preventing further misbehavior from likely offenders, zero tolerace policies can ideally reduce misbehavior. Broken-windows theory has faced some criticism [153], but in terms of explaining how school policy evolved it is more relevant that it has been believed than whether it is true. In *What Can Be Done About School Shootings?*, Borum et al. notes that roughly 75% of schools have zero tolerance policies for some offenses as of 2010 [24, p. 28].

4.1.1.2 The Crime Wave: Juvenile Superpredators and the War on Drugs

Crime rose quickly and steadily during the 1960s to 1980s in the U.S., and fear of crime has increased ever since the 1970s [65, p. 90]. While crime rates peaked in the early 1990s, people continued to think that crime was rising and fear continued long past that [72, p. *xi*]. In 1995 John DiIulio coined the idea of juvenile 'superpredators', or juvenile criminals who commit extreme levels of violent crime [87, p. 4] ("ultraviolence" from *A Clockwork Orange* [101]). As the idea of juvenile suprepredators was popularized and there was a rise in fear of crime generally, there was a turn toward stronger punishments in the criminal justice system, for crime in general and for juveniles specifically (who are now commonly tried as adults). In addition to leading to the prison boom, the high levels of fear of crime have meant that people were more afraid of violent juveniles, more doubtful of the idea of rehabilitation, and more willing to punish them.

Predating the idea of juvenile superpredators was the crack epidemic, in which crack cocaine appeared and spread through poor inner cities as a drug of choice. Richard Nixon and others essentially claimed that the drug use was the cause of a number of social problems in those communities (crime, violence, poverty, unemployment, etc.) and therefore greatly increased both penalties for drug use and rates of police enforcement for drug use [52]. In practice this was counterproductive, for two main reasons. First, while the negative effects of drug use are real, various social problems are more the cause than the result of drug use[3]; Barry Glassner notes in *The Culture of Fear* that crack was a "by-product of social and economic distress" but "became the *explanation* for that distress." [72, p. 136] This means that attacking drug use (or drug users) would do little to solve the problem. Second, the War on Drugs has been extremely harmful to the communities it reports to help. It spearheaded the prison boom (with people convicted of drug crimes now accounting for the majority of prisoners in the U.S.) that devastates poor communities of color in terms of community infrastructure, employment, and other measures.

As people thought that juveniles were more threatening than ever before, they subsequently

and Wilson did not advocate for zero tolerance policies per se. Rather the theory supported paying close attention to small details in policing before they escalate into critical problems; the theory was especially supportive of the notion of community policing. The link between broken window theory and zero tolerance was apparently developed by Mayor Rudolph Giuliani in his promotion of zero tolerance policing practices in New York City." [134, p. 136]. 'Broken-windows policing' and 'Zero tolerance policing' are often used more or less interchangeably, especially in reference to Giuliani and his Police Commissioner, William Bratton, who was essential in translating broken-windows theory into zero tolerance policy [111, p. 65].

[3]In his book *High Price*, neuroscientist Carl Hart challenges the idea that drugs are biologically addictive at all, claiming that the behaviors associated with addiction are caused by the social context surrounding the drug [5].

pushed for more aggressive, pre-emptive containment policies of potential juvenile threats [72, p. 74-5]. Zero tolerance was part of this, and the integration of police officers with schools was another. Police were on hand in case anything got dangerous, and because (people thought) delinquent juveniles were destined for a life of crime, they were given the power to arrest juveniles directly. The War on Drugs has been integrated with the school system, with a large number of drug arrests occurring directly in schools. Like superpredators, (poor minority) drug users are considered to be beyond hope, and best dealt with by expulsion and/or arrest so that they do not cause harm to the rest of the school. Thus, police have become a major part of the environment in many schools.

4.1.1.3 Surveillance Society

Since 9/11 the U.S. has increasingly been described as a "Surveillance Society", or a society that "function[s], in part, because of the extensive collection, recording, storage, analysis and application of information on individuals and groups in those societies as they go about their lives." [155] What this means is that systems of the society are highly dependent on monitoring the population and exercising power on the basis of these observations. Schools show signs of the surveillance society in video cameras, metal detectors, standardized testing (discussed below), and police officers in schools (among other things) [69, p. 72-3]. The School Crime Supplement to the National Crime Victimization Survey reports that around 69% of middle school and high school students report having police officers in their schools, while only 1% of schools reported police officers in 1975.[4] [117, p. 2]

The reasons for a growing surveillance society are complex.[5] First, a "culture of fear" [72], exacrebated by terrorist attacks on September 11th, has given the government wide latitude to employ surveillance technologies. Second, a rise in school shootings in recent years (in particular: shootings at Columbine High School in 1999, Virginia Polytechnic Institute in 2007, and Sandy Hook Elementary School in 2012) has led to a greater push for surveillance technologies in schools to protect the children.[6] Third, schools (like most organizations) are increasingly pro-data in the age of easy computing [69, p. 9], and surveillance technologies are read as signs of progress [115, p. 82]. For these reasons and others, schools have become increasingly dominated by surveillance, including police officers and other technologies that aid in the arrest of students.

[4]There is somewhat of a disagreement between the percentage of schools reporting use of police (40%) and the number of students reporting police in their schools (69%) [117, p. 2]. This could reflect dishonesty on the part of either the schools or the students (I personally would trust the students in that case), but it could also reflect a police presence being more common in large schools. Either statistic can be useful, although in terms of impact on students I think the student statistic of 69% is more relevant.

[5]Foucault essentially proposed the idea of surveillance society as coming much earlier (he puts the date at January 22, 1840, the opening of Mettray Prison) [60, p. 293]. While this trend is important in the appearance of zero tolerance policies, what is ultimately more relevant is the *expansion* of the role of surveillance in recent decades. Foucault's history stops around the time these major trends began (*Discipline and Punish* was published in 1977, and Foucault died in 1984), and so he does not really talk about the *recent* changes that led to zero tolerance policies.

[6]Columbine, Virginia Tech, and Sandy Hook actually had major surveillance and security systems at the time of the attack, and these proved totally unhelpful. Sandy Hook had enacted new security measures earlier that year, including locking the doors after 9:30am (which proved unhelpful when the shooter broke in by shooting through the glass) [89]. Virginia Tech updated its security system following its 2007 shooting, but exerienced another shooting in 2011. Two people were killed, despite the security system functioning without any glitches in technology and the high number of police officers on campus [31]. Columbine had extensive security footage and an armed police officer at the school already, but was unable to stop the shooting [160]. One author claims that the wave of increased security systems in schools was partially a result of the video footage of the shooting at Columbine [114, p. 3]; in other words, it happened only because Columbine had security cameras *and* they were ineffective in preventing the shooting.

4.1.1.4 No Child Left Behind and School Privatization

A less obvious part of surveillance societies is monitoring data like test scores. No Child Left Behind[7], the 2001 legislation that required regular standardized testing of schools for them to receive federal funding, put in place unprecedented systems for monitoring students, teachers, and schools. This has been expanded as part of a data-driven privatization movement, which has seen the rise of charter schools and Obama's *Race To The Top* program that effectively requires states to implement systems for assessing teachers to receive more federal funding [171]. Now, the progress of students is constantly monitored and used to judge teachers and schools, who are consequently punished or rewarded.

One major, insidious effect of this is that teachers are under more pressure than ever before to show positive results on tests. Testing takes away from class time, and many principles require teachers to teach test-preparation at the expense of the normal curriculum. As noted by educational historian Diane Ravitch, "...history, science, the arts, geography, even recess, were curtailed in many schools. Reading and mathematics were the only subjects that counted in calculating a school's adequate yearly progress [(which is based on standardized test results)], and even in these subjects, instruction gave way to intensive test preparation" [133, p. 107]. Even teachers who want to do more to develop students' abilities broadly and teach more deeply may not have that luxury, as their livelihoods are dependent on constantly improving test scores. As a result, they are likely to turn to more punitive measures for classroom management. Punishment is fast, easy, and often somewhat effective in the short term[8]; punishment that removes students from the classroom (through suspension, expulsion, or arrest) makes teachers lives much easier because they do not have to deal with the particularly difficult students. The end result is zero tolerance policies in the classroom, with police patrolling immediately outside the classroom.

It is worth noting, also, what types of schools exhibit the highest levels of punishment and are most likely to have police officers and zero tolerance. Broadly speaking, these affect poor and minority schools. In these schools, teachers have the fewest resources (in terms of teaching aides, school counselors, etc.), the larger class sizes, and the most difficult students (i.e., the students who come in with the lowest levels of formal education, English fluency, and experience in academic enrichment programs). Much like "mass incarceration" [108, p. 59], this trend is more notable for its concentration in already disadvantaged population than for its prevalence across different groups[9], and it would probably never reach the level it did if it affected different segments of the population equally. The continuing underinvestment in these schools, coupled with broader societal trends toward security, punishment, and pressures on teachers, causes teachers and schools to turn to more punitive measures for dealing with disruptive students.

[7]This aspect of surveillance was also a response to fear, specifically worries about the U.S. economy being passed by Asian nations like Japan, Russia, and China [139, p. *xv*], that had been at the forefront of political consciousness ever since the government report *A Nation at Risk* published in 1983 [170].

[8]In *Teaching Peace*, Beverly Title notes that "[s]chools may cling to the punishment paradigm not because they believe in it, but because it is fast and manageable. As an assistant principle once [said], 'I can suspend 10 students in the time it takes me to sit one Circle.' Sadly, true!"[165, p. 99]

[9]As noted earlier, the majority of schools do have some sort of zero tolerance policy, so in that sense it does affect most students. However, the exercise of zero tolerance policies disproportionately affects poor, disabled, and minority students; schools with large poor and minority populations have large numbers of suspensions, expulsions, and arrests, and within those schools racial minorities and the disabled are disproportionately targeted [172].

4.1.2 Effects of Zero-Tolerance and Policing

The turn toward more punitive policies has a two main effects. First, it is unlikely to help reform misbehaving students, and a review of theories of punishment in psychology show that it performs significantly worse than alternatives in changing behavior. Second, it disproportionately affects poor and minority populations, meaning these groups face more punishment and are perceived as more criminal. The result of these two effects together is the perpetuation of poverty and criminality. I discuss punishment in psychology before turning to the sociological effects of these policies on disadvantaged groups.

4.1.2.1 Theories of Punishment in Psychology

Theories of punishment in psychology fall broadly into four categories. Criminology has from its inception focused on punishment, starting with Beccaria's *On Crimes and Punishments* in 1756. Outside of criminology, the behaviorists were the first to study punishment in depth, and their scholarship is still perhaps the most developed on the subject. They were followed by motivation theorists, particularly those who focused on intrinsic motivation, and more recently by the theory of reintegrative shaming and the development of restorative justice in criminology. I discuss each of these four theories in turn, and their relevance to punishment in the classroom.

Beccaria Beccaria's major goal in his book *On Crimes and Punishments* was determining what made types of punishments were most effective. Beccaria thought that swiftness[10] and certainty of punishment were some of the most important factors in determining likelihood of re-offense, and that the severity of punishments was comparatively unimportant.[11] He thought that if punishments were related to the nature of the crimes committed, then punishment would be more effective, and that different levels of punishment for different crimes would reduce the rate of serious crimes.[12] Regarding moral certainty, Beccaria believed that crimes should be punished publicly to send a moral message[13], and that the death penalty increased crime because it reduced moral certainty. In summary, he said that "a punishment may not be an act of violence, of one, or of many, against a private member of society, it should be public, immediate, and necessary, the least possible in the case given, proportioned to the crime, and determined by the laws" [14, p. 38]. Each of these claims will be analyzed based on theories, studies, and evidence accumulated since Beccaria's writing was first published.

Beccaria claimed that punishment had to happen quickly for it to be most effective (see Thorndike's Law of Effect, later) and that the more sure subjects were of being punished, the more likely it would deter behavior. He also claimed that the severity of punishment was relatively unimportant in determining how likely people were to commit a crime. The first two claims have been broadly supported by evidence[14], and serve as a criticism of our current justice system (which

[10] "An immediate punishment is more useful." [14, p. 18]

[11] "Crimes are more effectually prevented by the certainty than the severity of punishment." [14, p. 23]

[12] "There is another excellent method of strengthening this important connection between the ideas of crime and punishment; that is, to make the punishment as analogous as possible to the nature of the crime, in order that the punishment may lead the mind to consider the crime in a different point of view from that in which it was placed by the flattering idea of promised advantages." [14, p. 18]

[13] "The public punishment, therefore, of small crimes will make a greater impression, and, by deterring men from the smaller, will effectually prevent the greater." [14, p. 18]

[14] For example, Mike Davis notes that swiftness and certainty of punishment deter crime significantly, and do so

is extremely slow, and for many serious crime is very uncertain). The second claim has also been broadly supported, with the caveat that it is true for punishments within a certain range but often untrue for sufficiently extreme punishments (e.g., the death penalty would probably deter talking in class more effectively than detention would). However, people are often reluctant to give out sufficiently harsh punishments, so in some sense there is a trade-off between severity and certainty of punishment.[15] Unwillingness to execute people for minor offenses made the Bloody Code ineffective in England 1688-1815 [42], and the same unwillingness (among many other factors) makes executing students for talking in class impractical as a punishment. This means that swift, certain punishment should be preferred to severe punishments, in school and elsewhere.

Beccaria believed that, when possible, having a punishment correspond to the crime was the most effective way to deter behavior, because it strengthened the association between crime and punishment.[16] Additionally, he thought that punishments should be proportional to crimes. To Beccaria, the absence of effective gradations of punishment based on offense meant that people would tend to choose the more severe crime over the less severe one (because there are no relative consequences), while effective gradations would cause them to choose the less severe one (e.g., if using a gun in a robbery significantly increases the criminal sentence, robbers wil tend to not use guns).

Finally, Beccaria talks about how punishments should be public to show the values of society and increase moral certainty, and how corporal punishments (especially the death penalty) could be counter-productive and reduce moral certainty (see evidence of violence increasing following wars). As long as criminal penalties are effective, evidence supports this claim, and public punishments have the additional effects of reaffimring society's values and helping people see the natural consequences of their actions.[17]

Beccaria's work, along with later supporting research, shows that swift and certain punishments are the most effective way to manage a classroom, and that severity is unhelpful or even counterproductive. His other main themes, that punishments should be connected to the crime and that it should be consistent with and re-affirm moral values, are equally important, but it is not immediately clear from Beccaria alone how punishments would fulfill these conditions.

Behaviorists The early behaviorists essentially saw punishment as the opposite of reward, and the ones who wrote about it largely said either that it was less effective than reward, or just not effective at all. Two theorists stand out here: Skinner noted that punishment also had a number

more effectively than increased severity of punishment [46], and the HOPE program in Hawaii has been very effective at reducing recidivism for people on probation by emphasizing "'swift and certain" punishment' [118].

[15]One way to get around this is building in a lot of procedural safeguards, as people will be more comfortable giving a harsh punishment if they feel the case has gone through a legal process designed to protect defendants. However, this means that there is a trade-off between swiftness and severity of punishment, and so the greater effectiveness of swiftness may more effectively deter behavior.

[16]Punishments are supposed to correspond to crime in a way that undoes the effects of the crime, in terms of harm and/or motivation. This means that zero tolerance policies that involve suspending students for truancy (or tardiness, to a lesser extent) are almost certainly counter-productive; this has not stopped them from being common [32].

[17]Durkheim believed that punishment as such was often unnecessary or counter-productive (saying that "a well disciplined classroom is one in which there is little punishment" [53, p. 160]) and that punishment, to the extent that it is used, should work to show social disapproval and to illustrate the natural consequences of behaviors [53, p. 167; 173]. Therefore, he thought that helping children see this was the most important part of their moral development (as contrasted with previous theories of natural consequences, in which the teacher ignores the child under the assumption that they will see the moral consequences themselves) [53, p. 170-3]. Durkheim's system also involves punishment corresponding to the crime much more effectively than Beccaria envisioned.

of other negative effects, while Guthrie noted that it could be effective if it involved the subject doing something else. Punishment was later reformulated with the idea of avoidance conditioning, in which it is not merely the reverse process of positive reinforcement but actually its own phenomenon that can cause escape and avoidance behaviors. Sidman further developed the paradigm and made more specific predictions, including that animals are willing to be punished more often in exchange for more control of their punishment, that they can detect difference in frequencies even when punishment is random, and that under certain circumstances animals may continue behaviors that cause punishments when other (punishment-free) options are available. Other behaviorists further developed the paradigm, mainly responding to Sidman.

Some of the early behaviorists did not address punishment at all in their theories, while most of the other behaviorists said it was less effective than reward in shaping behavior or that it was entirely ineffective. John Watson, father of behaviorism, did not address punishment in part because behaviorism was a new science (when he invented it), while Ivan Pavlov (the father of classical conditioning) did not look at punishment because his theory focused on stimuli outside of the organism's control. Clark Leonard Hull thought punishment was ineffective because he believed action was caused by innate drives [148, p. 249], although it is worth noting that his drive theory has fallen out of favor. Finally, Edward Thorndike started by saying punishment was less effective than reward, but later reformulated his "law of effect" to say that punishment was almost entirely ineffective, which is why he advocated effecting behavior change by uing rewards instead: "there is no evidence that [punishment]takes away strength from the psychological basis of the connection in any way comparable to the way in which a satisfying after-effect adds strength to it" [21, p. 90]. In general, early behaviorists provided little if any support for the idea that punishment is useful in learning.

Two of the early behaviorists had more nuanced opinions about punishment, though they still were not particularly favorable. B.F. Skinner, who developed the theory of operant conditioning, noted that punishment, in addition to being less effective than reward, could be harmful in a number of ways: "[i]n the long run, punishment, unlike reinforcement, works to the disadvantage of both the punished organism and the punishing agency" [149, p. 183]. Guthrie, who theorized that organisms pre-existing patterns of behavior are the best determinants of their current behavior, believed that punishment could be effective only if it caused an organism to change its behavior directly: "[p]unishment is effective only when it reconditions new responses to the cues for unwanted behavior" [78, p. 450]. These two theorists, while their view of punishment was a bit more complicated, did not differ from the rest of the early behaviorists: on the whole, they thought punishment was mostly ineffective.

Later behaviorists, such as Miller, Mowrer, Hernstein, and Sidman, developed the theory of punishment to include avoidance conditioning. Unlike previous theories of punishment, avoidance conditioning noted that punishment could be effective, at least in the limited context of causing people to avoid a certain behavior or thing [33, 112]. This leaves unaddressed theories like Guthrie's (that they need to learn other behaviors to replace the punished behavior), but supports the usefulness of punishment in a limited context. The main limit of avoidance conditioning is that people can acquire an avoidance response to any part of the environment, not necessarily the intended one. (For example, if a teacher punishes a child for hitting another child, that child may learn to avoid hitting other children, or may learn to avoid teachers, avoid school, avoid fighting in front of teachers, or any number of other (unhelpful, counter-productive) things.) Some other interesting results come out of avoidance conditioning: for example, one of Sidman's experiments

showed that rats are often willing to accept a greater level of punishment in exchange for greater control of that punishment [33], and this has been supported for humans by later studies. (This could explain "acting out" behavior: students can reliably be punished in predictable ways instead of risking being punished in different situations.) Broadly speaking, Sidman's avoidance conditioning and further work by behaviorists provided the basis for showing when punishment could be effective but provided little support for the idea that punishment was more effective than alternatives.

The clear take-away from this discussion is that punishment is ineffective, or rather that the punitive aspects of punishment are ineffective. There are a few specific circumstances in which punishment can be effective, and for the most part they involve the punishment directly changing people's behavior (as opposed to indirectly, through deterrence). Thus, a teacher's method of classroom management should be nonpunitive, and it should involve students doing desired behaviors.

Motivation Theorists In the behaviorist punishment experiments, many of the subjects would simply stop responding to shock or other punishment, and in general stop taking actions at all. In light of this, Martin Seligman theorized that these animals were suffering from learned helplessness, or the feeling that nothing they do makes a difference. Seligman claimed that learned helplessness was a result of organisms experiencing punishment that was unconnected to their actions (or at least appeared so), and that organisms experiencing punishment like this would develop learned helplessness [63, p. 127]. He noted how this theory is supported by studies of various human populations, including the elderly, SIDS, and people in the military [143]. Students who face high levels of punishment may have to choose between rebellion and learned helplessness, neither of which is a desirable option.

Deci and Ryan first theorized intrinsic motivation in 1985, in contrast to the earlier theories of behaviorism. Essentially, they noted that while behaviorist rewards (and punishments) were effective at controlling behavior as long as reinforcement was present, they were ineffective in the long term and often served to undermine behavior [47, *chapter 3*]. This is important because for subjects to acquire mastery in a given area, they have to be committed to practicing it for an extended period of time. Drawing on the theory of cognitive dissonance, Deci and Ryan noted that subjects who are 'rewarded' for a behavior may then feel that the behavior is only worthwhile if it will be rewarded, causing them to do it less often if reward is not present. Thus, it follows from this theory that what is normally considered 'reward' can function as punishment (in the technical Skinnerian sense) by reducing future behavior. A. S. Neill has used this effect at Summerhill School, and a common story in introductory psychology textbooks makes the same point.[18]

Alfie Kohn, in his book *Punished by Rewards*, notes the spread of "pop behaviorism", or a popular view of behaviorism that has become dominant in U.S. culture as people's understanding of behavior [98, p. 3]. This behaviorism is a simplified version of scientific behaviorism, less advanced even than the early behaviorists. Essentially, it is a simplified version of Thorndike's (original) law of effect: rewarded behavior will continue or increase, while punished behavior will stop or decrease. Kohn provides evidence of this failing to be the case across a wide variety of areas of life, noting that a behaviorist framework can undermine intrinsic motivation and thereby result in more long-term losses than gains (although it is effective in the short-term) [98, *chapters 7-9*]. He breaks from Deci and Ryan in noting how praise, normally considered to function like intrinsic

[18]Essentially Neill would reward a child for swearing by paying them money (a quarter). He paid the same amount for a few days and then slowly decreased how much he paid until he got to 0. During this time, students were motivated to swear by the money rather than by the pleasure of swearing, so that when he stopped paying them, they stopped swearing.[119, p. 24]

motivation, can function as extrinsic motivation (he notes this as happening much more broadly than Deci and Ryan: "we have four accounts of how praise may impede performance: it signals low ability, makes people feel pressured, invites a low-risk strategy to avoid failure, and reduces interest in the task itself" [98, p. 101].

Literature on motivation theory again supports the claim that punitive classroom management will be ineffective in causing deep behavioral change, and is likely to be counterproductive. Motivation theorists warn that controlling reward systems may suffer from some of the same problems (although probably to a lesser extent). Instead, classroom management should be student-focused. To the extent that the motivation and method of change can come from the student, it will be more effective, particularly in the long term.

Reintegrative Shaming and Restorative Justice John Braithwaite's *Crime, Shame and Reintegration* describes the theory of reintegrative shaming, which deals with punishment and integrates most of the leading theories and findings in criminology from the 1980s. It therefore gives most of the important changes in theory that have occurred since Beccaria, while also addressing the role of shame in punishment (and criminality). Using evidence from control theory, Braithwaite notes that more control of minor deviance from a well-respected and connected punisher will reduce or prevent serious lapses. Drawing on labeling theory, he says that stigmatizing labels will increase future criminality. Using social learning theory, Braithwaite supports Beccaria in saying that having public punishments will be most effective, and he supports the idea that punishment must be swift and certain because this leads to increased moral certainty (e.g., white-collar crime). Finally, Braithwaite says that the more punishment allows for (or intentionally involves) reintegration ceremonies, the more effective it will be in reducing future offenses.

Braithwaite predicts that strict control of minor deviance will reduce the incidence of greater deviance, provided the controller is well-respected and accepted as legitimate by the person being controlled [26, p. 55]. This is theoretically supported by the broken-windows theory (which is supported by Beccaria), but it makes a much weaker (and consequently better-supported) claim. The evidence used to promote broken-windows theory supports the idea that people's behavior may escalate if unchecked, as well as the idea that seeing other people's minor behavior fail to be punished can motivate other people to break the law. Since this requires a controlling authority accepted as legitimate, harsh punishment of low-level offenses is likely to be ineffective or counterproductive. Moreover, low-level deviance can be addressed without resorting to punishment, which may be seen as more legitimate. A disciplinary process in which students have more power to influence the outcome (as they do in restorative justice) will be more effective in general, whether for high-level or low-level deviance.

Braithwaite predicts that individuals who adopt criminal self-concepts will be more likely to continue to commit crimes (or to abandon the criminal self-concept), and therefore that punishment that encourages them to adopt these self-concepts will be counter-productive. Equivalently, he predicts that punishment that encourages them to move away from these self-concepts will be helpful in reducing future criminal behavior. Braithwaite notes that the evidence around labeling theory (the idea that labeling people as criminal increases their later likelihood of crime) is equivocal, largely because stigmatizing punishment can encourage *or* discourage the adoption of a criminal self-concept [26, p. 16-21]. Lonnie Athens, who examines self-concept directly, found that people's self-concept of themselves as violent or nonviolent matched the level of their violent criminal activity, which provides more support for Braithwaite's claim [11, p. 54-60]. This explains why reintegration

(or lack of dis-integration) can be useful in punishment; students who remain integrated with a classroom are less likely to adopt criminal self-concepts.[19]

To a large extent Braithwaite reiterates themes from Beccaria about how punishment should be related to the infraction and how it must be seen as legitimate to be effective. Braithwaite more explicitly talks about how offenders must continue to be integrated with current communities, and brought back into them whenever they leave. In addition, they should be individually empowered to make the necessary changes to their behavior in ways appropriate to them individually.

Negative Effects of Suspension, etc.

4.1.2.2 Perpetuation of Poverty and Criminality

In *The New Jm Crow*, Michelle Alexander talks about how prsons have become part of the dominant system for controlling and oppressing poor blacks in the U.S. She effectively describes how racism operates through the War on Drugs to marginalize black men during the arrest, conviction, and re-entry phases of the criminal justice system: "The War on Drugs is the vehicle through which extraordinary numbers of black men are forced into the cage. This entrapment occurs in three distinct phases..." [3, p. 180]. By the end of this process, black ex-felons face all of the types of discrimination that were present in Jim Crow before the Civil Rights Movement, except this now occurs through nominally race-neutral policies [3, p. 181]. Black men affected who go through the criminal justice system are likely to remain in poverty, and lack of legitimate job opportunities after prison mean they are likely to return to prison also. Thus, the prison system primarily affects poor black men, and makes them more likely to be poor black criminals in the future.

Alexander notes that, while the justice system and drug wars have always been racially biased, this War on Drugs is qualitatively different because of its reach. She says that "the War on Drugs has given birth to a system of mass incarceration that governs not just a small fraction of a racial or ethnic minority [as previous drug wars did] but entire communities of color. In ghetto communities, nearly everyone is either directly or indirectly subject to te new caste system...This time, the drug war *is* the system of control" [3, p. 183]. Loïc Wacquant makes a similar claim in *From Slavery to Mass Incarceration*, noting four 'peculiar institutions' used to control blacks in U.S. history: slavery, Jim Crow, Northern ghettos, and the hyperghetto and prison system [174]. Thus, the War on Drugs and the prison system are important because they are the primary forces working to contain and control blacks in the U.S. This is relevant because imprisonment and the War on Drugs directly interface with discipline in schools, but also serves an an illustrative example: dropping out of school (including suspensions, expulsions,and arrests) has many of the same negative effects that perpetuate crime and poverty.

As mentioned above, zero tolerance policies disproportionately affect poor and minority populations. This happens in two ways. First, they are disproportionately implemented in poor and minority schools, which often face the highest pressures to increase test scores. Second, within schools that have zero tolerance policies, minority students are disproportionately punished (by teachers, administrators, and police officers). We have seen from reviewing the literature in psychology that this does not help behavior, and in fact tends to make it worse. Thus, these populations are disproportionately punished, and they suffer for it.

[19]Braithwaite notes that the "association between poor school performance and delinquency is universally strong, as is the relationship between not liking school or being weakly attached to school and delinquency." [26, p. 29]

There are additional negative effects of this punishment, beyond breeding recalcitrance and worse behavior. First, it perpetuates stereotypes of poor and minority children as criminals. Because they are disproportionately punished (and because they continue to act out, since the punishment is ineffective), people assume that they are worse-behaved than other students. Having police officers patrol the hallways, and in some cases even arrest these students, perpetuates this idea of criminality further. Thus, every school that displays discriminatory punitive policies encourages other schools to do so also, because people think that certain populations are more criminal than other populations.

Most importantly, this perpetuates poverty and criminality among poor, crime-prone populations. The effects of expulsion and imprisonment on economic mobility are well-documented; they lead to higher rates of unemployment, reduced wages, and much lower mobility [176, p 112]. Similarly for criminality: for populations who are already more likely to commit crimes (those with behavioral problems and lower levels of schooling), removing them from school in any capacity is likely to only increase criminal dispositions. When they are arrested and imprisoned, they are especially likely in the future to commit crimes because they have become part of that system [60, p. 298-302]. Thus, overly punitive measures in schools, when directed at poor populations with criminal tendencies, make them more likely to be poor and more likely to break the law in the future. Poverty continues across generations even more so than it does already, and people on the path to becoming criminals are pushed further into criminality. In the long run this can lead to an intensification of policing and zero tolerance in schools, making the cycle even more harmful.

Conclusion: Relevance to Proposal

Some major themes emerge from this literature review about what needs to be changed in education for effective, nondiscriminatory classroom management. Changes in national policy that decriminalized adolescence, reduced NCLB and similar requirements, and created nonpunitive alternatives to prison would be ideal, but there are changes that can occur at the level of individual schools to combat this problem also. This proposal aims to do exactly that, with three main focuses: keeping students in the classroom, rehabilitating problem students, and making things easier for teachers.

Keep Students in the Classroom! The first theme from here is the importance of dealing with things in the classroom. Losing a day of class for suspension isolates and stigmatizes students, and interferes with their learning. Expulsion and arrest are even more dramatic, in dramatically curtailing students' future opportunities in life. In part because of this, many schools have recently implemented in-school suspension programs. For an in-school suspension (ISS), a student is removed from the normal classroom and placed in an ISS center at the school, where they are subject to various behavior-change strategies (and possibly homework) [19, p. 2]. ISS has shown success in some cases, and ineffectiveness in others. In a major study looking at experimental ISS programs in 20 matched school districts, researcher Larry Leapley found that "all of the experimental schools noted a significant change for the better" in terms of reducing violence [19, p. 8]. This program was effective because of a trained, effective teacher staffing the ISS room. Some other schools that have implemented ISS programs have had less success because of lack of resources, students using ISS to get out of class, and teachers using it as a way to remove disruptive students from their classrooms without addressing the root problem (much like they do with out-of-school suspension) [7, 19].

Essentially, in-school suspension can be effective when well-run, but in practice it often suffers from many of the problems associated with out-of-school suspension. When there are not resources

to support it, ISS is ineffective. Similarly, when ISS is used as a way to remove students from class (either by teachers or by bored students), its use is inappropriate and ineffective. Keeping students in the classroom avoids all of these concerns. Teachers who are able to deal with problems in the classroom will be more effective at reducing misbehavior than under-resourced ISS rooms, and they will also be able to teach students academic content more effectively. Students who will not be suspended cannot use misbehavior as a way to get out of the classroom, and teachers cannot use suspension as a quick (but ineffective) fix for student misbehavior that fails to address the deeper problems.

Another advantage of avoiding suspensions of all types is that they are highly discriminatory, while classroom management has less potential to discriminate in important ways. The reason discrimination in punishment is so important is because of the negative effects that go with removing students from the classroom. By learning how to deal with disruptive behaviors without excluding students, disruptive students will have much better life chances. Because exclusionary punishment is an easy short-term solution, teachers frequently opt to remove students from the classroom in place of finding and addressing the root causes of their misbehavior. If teachers (and administrators) do not have that option, they are more likely to address the important root causes of negative behaviors.

Rehabilitate Misbehaving Students As noted above, behaviorists are fairly unanimous in saying that punishment is ineffective as a means of deterring behavior. Nor is punishment effective for building motivation; in fact, punishment is usually counterproductive in motivating students to succeed or behave in the classroom generally. Reintegrative shaming helps students take more responsibility and denounce their own problematic behaviors without hurting their self-concept. Looking beyond just reintegrative shaming, taking responsibility for one's actons is a major part of how resotrative justice works generally. In restorative justice, the offender looks at the harm caused by their behavior and how they can work to repair harm already done and prevent additional harm in the future [185, p. 20]. Restorative justice programs have proven effective at reducing recidivism statistically, and at encouraging qualitative personal growth [61, 169]. To the extent that classroom discipline can focus on prevention and repairing harms rather than punishment, students will be less likely to continue their negative behaviors and the classroom environment will be more positive.

By focusing on student development and finding solutions rather than merely punishing, this curriculum can work to make students more functional, better-behaved people who have the tools and willingness to deal with problems in school. Classroom structures for dealing with conflict help students while there and, ideally, in the future as well. Giving students responsibility in resolving conflicts, rather than imposing solutions on them, empowers them to self-actualize. Moreover, giving students responsibility allows the teachers to create high expectations for students, while also allowing students to make mistakes. The policy I propose is 'zero tolerance' in the sense that deviance *that students do not then work to improve* is not tolerated, but it is forgiving and offers second chances. Because the program being proposed (if successful) would show that these students can be rehabilitated, it also combats images of them as criminals and helps them move beyond past misbehaviors (the opposite of a criminal record).

Make Teachers' Jobs Easier The two biggest educational policies of the millenium have probably been No Child Left Behind (NCLB) and Race To The Top (RTTT). NCLB as a policy is

extremely punitive to schools that fail to make Adequate Yearly Progress[20], in that schools are frequently required to close, be taken over by the government, or otherwise restructure. It is also is coercive, requiring schools to implement heavy standardized testing regimes in order to receive funding. RTTT allows states or districts to win money by implmenting certain changes (like merit pay, faculty assessment, and no limits of charter schools) and is therefore slightly less coercive and punitive, but since schools are consistently underfunded it still pushes them to adopt possibly destructive reforms they otherwise wouldn't [133, p. 218]. While individual schools could beenfit from these policies, NCLB and RTTT have only managed to affect national education policy significantly because of their coercive nature; otherwise, they would not be adopted by most schools and districts. Resistence from districts, teachers, and students demonstrates this.

Because schools are pressured to adopt NCLB and RTTT against their will, teachers' jobs become more difficult and their teaching is disrupted. Time spent testing takes away from time in the classroom, and pressures from testing narrow the curriculum. At the same time, school closure or restructuring and the firing of teachers because of low assessments undermine stability in the system, further complicating teachers' jobs and threatening the public school system as a whole. A policy that is not forced upon teachers will not suffer these problems, nor will it face significant resistence that undermines it effectiveness. Instead, teachers who find it helpful will choose to adopt it, and thereby improve their classrooms.

As noted above, a large part of the shift toward punishment in schools has resulted from increasing pressures on teachers. Any program that creates more work for teachers is unlikely to suceed, however well-designed otherwise, both because such a program pushes teacher toward punishment naturally and because it is likely to be resisted by the teachers incharge of its implementation. To be successful, therefore, a policy should reduce teachers' burden and make managing the classroom easier for them rather than more difficult. Thus, my proposed policy should be easy for teachers to implement and should improve behavior in the classroom sufficiently that they find it worthwhile to do so.

The policy below attempts to fulfill these three goals. It has three parts, ideally implemented together but functional as separate entities so a teacher may choose to implement only part of the policy. It keeps students in the classroom and it focuses on helping them learn to prevent or solve problems in the future. The hope is that it will also make teachers' jobs easier in terms of classroom management.

4.2 Proposal

4.2.1 In-Class Management

This sections describes a process for teachers to resolve conflicts in the classroom. This is organized into three parts: preparation, addressing the conflict in class, and follow-up.

Planning First, teachers should plan every class with a short back-up activity, with two important characteristics: it relates broadly to the topic being studied, and it can be employed at any

[20]Schools have to make AYP is every subgroup of their school population to avoid this punishment. It is common for an otherwise very successful school to fail in one category, typically special education students, and therefore not meet AYP [133, p. 100].

time. Another important part about this activity is that students are able to do it more or less independently of the teacher, either individually or in groups. Examples of this include independent reading time, riddles/puzzles (in math, language, etc.), think-pair-share activities (with the think-pair parts), or perhaps even mini-presentations by students (more ambitious, but could be very beneficial to students). If a student (or students) is disrupting the class and the teacher needs to deal with them more individually, the teacher can assign the rest of the class to this short activity and talk with the student.

Short-term Dispute Resolution If the issue is sufficiently non-serious, the teacher meets with the student and creates a temporary fix to the problem. Issues could involve hunger, crying, not sitting still/making noise, fights with other students, or other things. In most of these cases, the teacher sits down with the student and has them explain the problem. THe teacher explains in a clear (and nonjudgemental) way how the behavior is disruptive to the class and helps them find (at least short-term) solution (e.g., giving food, nonpunitive time-out, etc.). In the special case of two students fighting, the teacher is encouraged not to intervene or take sides and let the children quietly resolve it on their own. Teacher may neutrally/multilaterally mediate if they are confident in their ability to do so.

Circles, Conferences, etc. For serious or frequent problems, teacher may find it useful to meet with student outside of class to work on longer term solutions. This could be mediation, conferencing, or circle processes (depending on problem), and could occur during lunch, recess, or after or before school. These processes are detailed below:

Mediation Mediation can be quite difficult for someone without training, so when possible it's usually best to let students who are fighting resolve disputes on their own. This is why Carl Rogers's mediation rule is taught below, both to give students a structure to resolve disputes on their own and to help teachers inexperienced in mediation by limiting their role (to enforcing this rule). Mediation is included here as an option because in the case of an ongoing problem, a teacher may need to create tme and space after class for disputants to resolve the issue.

Conferencing Parent-teacher conferences are a well-established tradition in schools: essentially the intent is to involve parents more directly so that they are aware of the effects their child is having on the classroom environemnt. Pressure from parents can make children less likely to misbehave, as well as helping the teacher understand the situation better. Parents can also come up with effective solutions and help implement them (for example, by making sure the child does their homework, gets to class on time, etc.). It is important that conferences are not stigmatizing to the child; instead, they should focus on the negative effects of their behavior but actively work to bring the chld back into the classroom.

Circle For very serious or ongoing problems, a circle may be necessary. Circles are quite resource intensive (in terms of time they take up and number of people involved), but because they involve all the relevant stakeholders and go into depth about the process, they are best equipped to deal with serious problems [129, p. 153-167]. Someone other than the teacher should be the facilitator, so that the teacher can be a more effective participant as an affected party.

4.2.2 Conflict Resolution Pedagogy

This section lays out a basic curriculum for students in conflict resolution, with additional levels of conflict resolution for teachers who feel it is helpful to teach more. This can be taught in individual classrooms or at the level of the entire school. In any case, an important part of this program is that the teacher(s) continue to use these conflict resolution ideas throughout the year so that they contribute to the classroom on an ongoing basis. This curriculum (which could be taught within the span of a class) includes the most basic, easy-to-implement aspects of conflict resolution that could still contribute positively to a classroom. Again, they can be taught quickly but must be integrated with management of the classroom in the future.

Barriers to Communication Robert Bolton's *People Skills* describes 13 barriers to communication [22, p. 17]. The first twelve are below, and fall into three categories:

- Judging
    1. Criticizing
    2. Name-Calling
    3. Diagnosing
    4. Praising Evaluatively
- Sending Solutions
    1. Ordering
    2. Threatening
    3. Moralizing
    4. Excessive/Inappropriate Questioning
    5. Advising
- Avoiding the Other's Concerns
    1. Diverting
    2. Logical Argument
    3. Reassuring

These barriers can be presented to students initially and then talked about throughout the class (as something people in novels do, or talking about a past conflict later). Importantly, the Thirteenth barrier to communication is pointing out that someone else is using a barrier to communication; mentioning it *during* an argument is usually counterproductive. (On the other hand, if a neutral third party, such as a mediator, or a friend points it out during an argument, it may be helpful.) However, discussing specific conflicts later can be helpful for having students reflect and learn (especially during problem-solving). Additionally, the teacher can hang a poster, so students are reminded of and thinking about the barriers to communication.

3-Part Assertion Messages Three-part assertion messages are a way to communicate that it clear, specific, and less aggressive. People describe a specific behavior the other person is doing, their feelings that result, and why they feel that way [22, p. 143]. (For example, "When you don't clean the counter after baking, I feel very annoyed, because it makes more work for me.") Pointing to a specific action expressed in a nonjudgmental way helps people understand exactly what it causing a problem, and the other two parts explain why without being accusatory. This is how teachers should communicate problems with students and how they can coach them to address problems with other students. Teachers can remind students to communicate this way, which de-escalates and helps them calm down.

Carl Rogers Mediation Rule This rule, employed by humanistic psychologist Carl Rogers, states that in an argument, "[e]ach person can speak up for himself only *after* he has first restated the ideas and feeligs of the previous speaker accurately, and to that speaker's satisfaction." [135, p. 332] This rule explicitly requires disputants to understand others' perspectives before they can respond to them, and empowers them to reiterate if they feel other people do not understand their perspective. In mediation it is an incredibly useful tool for teachers, and probably one of the only healthy ways for them to get involved in a mediation. Even if they do not get involved they can encourage students to do this when they are arguing with each other, and it may be useful for in-class discussions and debates as well.

4.2.3 Teacher Training

Teachers should first of all learn whatever conflict resolution techniques are being taught to students, meaning the curriculum in 2.2. In addition, there are some skills that would be especially beneficial to teachers (and perhaps less so to students): reflective listening, handling pushback, and a conflict resolution model.

Reflective Listening Reflective listening, the ability to listen to someone else and repeat back their feelings and the cntent of what they said, is a skill suseful to everyone but especially important to teachers [22, *chapter 4*]. While this is a vital skill to conflict resolution, it is one that requires more effort on the part of the listener, and thus may be more difficult to implement with students than the Carl Rogers rule, which is clearly limited in time and context. Teachers benefit because it makes their lives much easier. As a teacher I would talk about this with more interested students, but I am not sure it would be very helpful to try to teach and enforce at the level of the whole class. At the very least, it depends how much effort the teacher wants to put in to improving students' communication skills (though this would be rewarding to everyone if implemented).

Handling Push-Back This extends the skills from three-part assertion messages, because sometimes they can (initially) backfire if people are not accustomed to criticism. This is fairly straightforward to deal with, but can be more emotionally difficult because of the feeling of being attacked. Essentially, what someone does is response to push-back is to listen reflectively to the other person. Doing this a few times will show the other person that they are being listened to, at which point the reflective listener reasserts their request, making sure to use the exact same wording as before (or as close as possible) [22, *chapter 10*]. This process may have to be repeated a few times, but the push-back normally decreases each time until the other person is willing to listen to the assertive statement.

Conflict Resolution This follows fairly directly from most of the skills before. In the basic phase of conflict resolution, the party using the method listens until they are able to experience the other side, or feel and understand it as well as possible (this is confirmed with reflective listening). They then asserts their views, needs, and feelings with three-part assertion messages. Oftentimes this will resolve conflicts directly, when a solution is easily in sight or one side just needed to vent and be heard [22, *chapter 13*]. If not, use the collaborative problem solving method in the next paragraph. Regardless of whether that is used, the system works fairly well when only one person is trying to use it, making it an ideal tool for a teacher who may have conflicts with uncooperative students, parents, peers, and administrators. It is best if both people use it, though, so this would also be a helpful skill to teach students.

Collaborative Problem Solving: The first part of this is defining the problem in terms of both people's needs rather than in terms of predetermined solutions. If someone has a solution in mind, finding the need satisfied by that solution allows for the possibility of other solutions, including one that may satisfy both people's needs. The next two steps are brainstorming possible solutions and then selecting the solution that best meets both people's needs (after trying to forsee possible consequences of that solution). Finally, people should plan and explicitly decide who is going to implement each part of this solution, and then ultimately these parties must implement it [22, *chapter 14*]. If possible, both people reflect on the process later to make sure the conflict was actually resolved and find out which parts of the process were effective.

4.3 Implementation Concerns

As much as possible this policy has been designed to be easy and attractive to implement. The goal is to make teachers' jobs easier, so that this should be implemented by teachers in part for their own benefit. Materials supporting each part, while not included here, will be produced if I get my grant and supplied to schools attempting to implement the policy in order to make the process even easier.

One important note is that this is unlikely to be effective if imposed as a top-down policy. State and federal governments, or even district leadership lacking in massive teacher buy-in, would not be able to impose this policy on unwilling teachers and schools effectively. (Even the professional development piece, which they could institute as a requirement for teacher licensure, would be unlikely to make a difference if teachers did not wish to use the techniques it taught.) The three parts of this policy are designed to be modular so that a school or teacher may choose to adopt some, all, or none of them as they find helpful. While the assessment of this policy requires schools willing to implement all of these parts, if the program proves successful, materials can be made available to teachers only interested in implemented some of them.

4.4 Assessment

Classroom management is not a variable that can be directly measured, but its effects can be inferred from a number of related measures such as number of disciplinary write-ups, qualitative self-reports of classroom environment, and amount of time spent on pedagogy versus classroom management. Both because this program is designed to benefit low-income schools with classroom management problems and because looking at classrooms with high levels of disciplinary problems

allows the sample size to be smaller, this program shall be tested in low-income schools with high levels of disciplinary write-ups.

To start this study, three to seven low-income public schools with high levels of disciplinary infractions are selected randomly and matched with other, demographically similar schools. (Depending on costs each school could ideally be matched with five to ten such schools; at a minimum each school must be matched with one other school, and the number of experimental schools must be higher if each is paired with only one school.) Before schools are chosen, different schools will be asked whether they would be interested in participating, then some subset of those will be selected (this prevent selection bias). After a period of three to five years of full implementation of the program (i.e., after teachers have been trained, taught conflict resolution skills to students, and been using the classroom management procedures for three to five years), data analysis will compare the schools to what they were like before the program and how to fair compared to similar schools. If possible, longitudinal data about students will be compared ten to fifteen years after the start of the program. Data are measured in the three categories: disciplinary data, other school measures, and longitudinal data.

4.4.1 Disciplinary Data

Detentions Detentions are one of the most important punishments to measure, because they are the most common type of punishment (which means more statistical significance) and are frequently the default punishment for teachers. A detention is not entirely different from an out-of-class conference, circle, or mediation, except that a detention is more directly punitive. As a researcher I would probably count any of these non-punitive class meetings as equivalent to a detention for measuring disciplinary problems (even though it seems preferable, it is presumably a result of problems in the classroom). Thus, detentions and conferences, mediation, circles before the policy is implemented and in schools without the policy are compared as a group to schools with the policy (after implementation).

Suspensions and Expulsions Suspensions and expulsions are a stronger form of punishment than detentions, and are for most purposes worse because they deprive the students of class time and because their use often reflects teachers moving or hiding, rather than dealing with, problems in the classroom. A significant reduction in suspensions or expulsions, even if it means a comparable rise in detentions, means that the policy is successful in helping teachers deal with conflict in a reintegrative way. Many public schools have quite high rates of suspension and expulsion, meaning a significant effect is likely to be reflected in the data.

Arrests As police officers have acquired a significant presence in many schools, arrests have become common enough to measure and analyze statistically. This program could reduce arrests for two primary reasons: it can cause the schools to intentionally use police officers less because they discipline students in the classroom, or students can avoid arrest by behaving better. In the first case, the school has found other ways to respond to behavioral problems at the school, which means the program is succeeding. In the second case, the program promotes better behavior among students, which is also a sign of policy success. While a qualitative examination of the school may be needed to determine the exact cause and mechanics of a reduction in arrests, either way indicates success in the policy.

Other The school may keep disciplinary records of other types, such as the number of times students are sent to the office during class (without receiving a detention or being suspended, expelled, or arrested). A school may have teachers write reports whenever a student misbehaves in class so that they can address this more proactively. When these or related measurements are present in schools, they should also be measured to see how effective the program is. In an example like the second one, where a teacher records student misbehaviors without there being punitive consequences, this should continue to be measured (so that a reduction in write-ups is not an artifact of (some) teachers no longer recording, but rather a difference in behavior). On the other hand, if teachers no longer send students to the office because they can deal with the problem in class, this means that policy is succeeding in at least keeping students in the classroom and it should be noted as such.

4.4.2 Other School Measures

Graduation (and Grade Completion) This is probably the most important non-disciplinary measurement, both because it probably reflects a real difference and because one of the major goals is to help students stay in school. In addition to actual graduation, the number of grades completed by students can be measured (e.g., a school that has students finish sophomore year before dropping out is preferable to one where students drop out before completing freshman year).

GPA GPA is one of the most stable measurements of academic achievement, as well as one of the best predictors of future outcomes. If this disciplinary program is effective in building a better classroom environment, it may result in students having higher GPAs. There are two major concerns about measuring this as a variable, however. First, if the school grades students on a curve (formally or unconsciously), there will be no difference in GPAs after this program is implemented; gains of one student necessarily come at the expense of another student(s). Second, if this program reduces expulsions and arrests for high-risk students, GPAs may fall because these students are likely to have worse grades. Thus, GPA is potentially useful as a measure, but only if the school does not use a grading curve and it can be compared with another school that lost the same number of students. This is a standard measurement worth including, but graduation and similar things take precedence.

SAT/ACT, or other Standardized Tests Standardized test scores have most of the same strengths as GPAs (although to a lesser extent); they are relatively consistent over time and predict a variety of future outcomes. The are better than GPAs in that there is no concern of grading students on a curve; unless the entire nation were to implement this program, a significant change in education would likely be reflected in test scores. However, there is the same concern about low-performing students staying in school and therefore pulling down scores.

Number of Honors Courses This is another academic measure, but it is relatively independent of other students (compared with bell curve-based grading). Students may become more motivated, be placed in higher tracks because they can handle discipline better, or improve enough academically that honors classes become appropriate. Again, this may be influenced by students (not) leaving the school: if the program is successful in preventing low-performing students (who are behavioral problems) from dropping out, the average number of honors classes taken may decrease even if individual students become more likely to take honors classes.

In assessing this program, disciplinary measures are primary (starting with the most extreme and working toward detentions), followed by graduation, followed by academic measures. More than anything this program is focused on reducing behavioral problems in the classroom and keeping people in school (and in classrooms). Academic achievement is both less important and more difficult to measure reliably, and therefore works more to add information or function as a tie-breaker than primarily assess the effectiveness of this program.

4.4.3 Longitudinal Data

Research in psychology is usually lacking in longitudinal studies, mainly because of the high cost associated with conducting them. Because this proposal is closer to a pilot program there will not be an attempt from the outset to collect longitudinal data. On the other hand, if effects prove to be exceptionally positive, this success would hopefully inspire a longitudinal study. Alternately, a researcher could follow up with the original students from this study. Relevant variables would be detentions, suspensions, and expulsions in future classes, arrests, college graduation, and GPA, among other things.

4.5 Conclusion

The rise of zero tolerance policies in recent years has hindered student learning and is perpetuating cycles of poverty and criminality. This policy aims to help individual teachers with their classroom management, both by making their jobs easier and therefore improving their pedagogy and by giving them another alternative to harsh, zero tolerance sanctions in the classroom. The classroom management approach outlined here hopes to also help students, by allowing them to stay integrated with the classroom and learn more and by helping them take responsibility for their own behavioral change. If this policy is helpful to teachers and schools, other schools will take notice and be interested in using these methods. In this way, this policy or similar alternatives to punishment may spread through schools and help combat zero tolerance policies and continuing poverty and criminality.

Part III

Independent Components

Chapter 5

Painting by Numbers: A Two-Part Examination in Statistics and Data Visualization

NSC (1729)
Statistics and Data Visualization
Fall 2013-Spring 2014, Exam Syllabus and Rubric

Student: Daniel Kalla

Exam Writer and Evaluator: Matt Ollis

Overview

This exam will look at my skills in statistics and data visualization, with a slightly stronger focus on statistics. It is intended to encompass most of my work in both fields, and allow me to demonstrate my skills in them. The exam will have two parts: a timed, written portion completed in a single sitting, and a take home portion completed in the span of a week. The purpose of this is that the ability to do statistics and understand data visualization is sometime useful, although because the ability to do them well overall is more important generally is more important the take-home portion will comprise a much large percentage of the exam in terms of grading. To keep the exam manageable in scope, not all possible topics will be part of the exam, with the understanding that I will prepare for all of them and that my skills in the areas tested are probably comparable to my skills in the areas not included. In statistics, potential topics include the regression model (including multiple regression, time series, and instrumental variables) to show a foundation in the dominant types of statistics, and nonparametric statistics and issues surrounding hidden populations to show a facility in less common areas of statistics, and various qualitative issues important in interpreting and using statistics, particularly in the social sciences. The exam may also include a research section, in which I demonstrate my ability to find relevant data, and a section that requires me to use a new type of statistics, showing that Im able to learn additional skills in statistics as needed. In data visualization, topics include different aspects of the theory of data visualization, as well as my ability to create and to critique different data visualizations. This exam is prepared by my professor, advisor, and former employer (for math tutoring), Matt Ollis. Below are learning objectives, major foundational texts in my course of study, specific percentages of grading, and areas of curriculum covered in more detail.

Objectives

By completing this exam, I should show that I am able to:

1. Hypothesis-test and calculate statistics with a regression model
2. Find statistical information through research
3. Discuss qualitative issues in using and interpreting statistics
4. Research and analyze non-normal or hidden population
5. Learn and apply new types of statistical analysis
6. Explain and apply the theory of data visualization

8. Critique other visualizations (and offer suggestions for improvement)

Grading

The method of grading this exam will be up to Matt's disgression, with the exception of the relative weights of the two parts, below:

Sit-down Paper Exam – 30%
Week-long Take-home Exam – 70%

Foundational Texts

- Econometrics: A Modern Approach (Wooldridge)
- Respondent-Driven Sampling (Heckathorn; Knaggs)
- Non-parametric Statistics (Higgins)
- How to Lie with Statistics (Huff)
- The Black Swan (Taleb)
- Various Articles about Statistical Significance (Deirdre)
- Why Most Published Research Findings are False (Ioannidis)
- The Functional Art (Cairo)
- Show Me the Numbers (Few)
- Visual Explainations and Envisioning Information (Tufte)

1 Curriculum Outline – Statistics

1.1 Econometrics

- Nature of Econometrics and Economic Data
- Simple and Multiple Regression: Estimation and Inference
- Asymptotics, Heteroskedasticity and Further Issues
- Binary Variables, Specification, and Data Issues
- Regression Analysis in Time Series
- Serial Correlation, Heteroskedasticity, and Further Issues with Time Series
- IV Analysis and TSLS

1.2 Research

This section involves being given a research question and then asked to research it (online or at Marlboro). My submitted work here includes a description of my process, as well as an answer to the research question (if possible).

1.3 Hidden Populations and Nonparametric Statistics

Hidden Populations

- Sensitive Survey Questions
- Interviewing Strategies
- Respondent-Driven Sampling

Nonparametric Statistics

- When to use nonparametric statistics
- Basic Inference and Comparisons of 2 or more samples (including paired and block samples)
- Correlational nonparametric statistics

1.4 Qualitative Issues in Statistics

- Statistical (vs. Practical) Significance
- Publication Bias
- Proxy Variables
- Causation and Correlation
- What Statistics Measure
- Black Swan Events

1.5 New Statistics

In this section, I am given a type of statistical analysis different from those earlier in the test and expected to do some research and perform a statistical analysis.

2 Curriculum Outline – Data Visualization

2.1 Theory of Data Visualization

Categorization of this section and the following is a bit arbitrary, but topics are divided to the best of my ability

- Why Visualize Data? Advantages of Visualization, and Types of Visualizations
- Art vs. Functionality, and the Complexity Challenge
- Cognition and Perception of Data Visualization

- Types of Graphs and their Purposes
- Interactive Graphics
- Roles and Types of Tables and Graphs
- Small Multiples and Least Easily Perceivable Difference
- Non-graph, non-table Data Visualization

2.2 Creating Infographics

- Process of Creating Infographics
- Design of Tables and Graphs
- Component-Level Design of Graphs
- Displaying Many Variables at once and Highlighting Data
- Using R

2.3 Critiquing other Data Visualizations

This will involve looking at examples of data visualizations and assessing what works well, what doesn't work well, how I would change it, and what I would experiment with (these don't necessarily overlap).

Statistics and Data Visualization: Final Exam 1

Answer all questions. You may use the statistical software `R`, but not books or notes (you may use sources for help with `R` syntax; if you do so say exactly what you consulted). You have 3 hours to complete the exam. Express your arguments clearly and carefully and state any additional assumptions that are required to answer the questions. How you reach an answer is at least as important as the answer itself.

Data sets are either built into `R` or available via the `openintro` package. In the former case they are immediately available in any `R` session as a data frame and documented in the help section (see, for example, the documentation fro the iris data set used in question 1 by typing "?iris" and the data set itself by just typing "iris"). In the latter case, you will need to install the package ("install.packages(openintro)") and then call it ("library(openintro")). You can now access the data sets there as you would the built-in ones.

Good luck!

1. Access the iris data set built into `R`. Is sepal length correlated with petal length? If so, does this relationship vary with species? Draw at least one graph to include in your explanation.

2. You intend to conduct a study of drug use among graduates of Marlboro College, specifically to investigate whether Marlboro graduates differ from similar populations. Discuss at least three significant data collection issues that you will need to address and how you might address them.

3. Consider the following interactive visualization of the interdisciplinary nature of plans at Marlboro:

`http://www.marlboro.edu/test/plan/16/`

Comment on its strengths and weaknesses. What might you do differently if you were designing it?

Same question for this interactive visualization of smoking data:

`http://viz.healthmetricsandevaluation.org/tobacco/`

4. Access the oscar data set in the `openintro` package. Collect some summary statistics and create some graphs that illustrate potentially interesting aspects of the data.

5. Consider the following simple regression model for salary (*wage*: dollars/year) and education (*educ*: years of education) to be used on a data set for men in the US aged 25–44:

$$\log(wage) = \beta_0 + \beta_1 educ + \epsilon$$

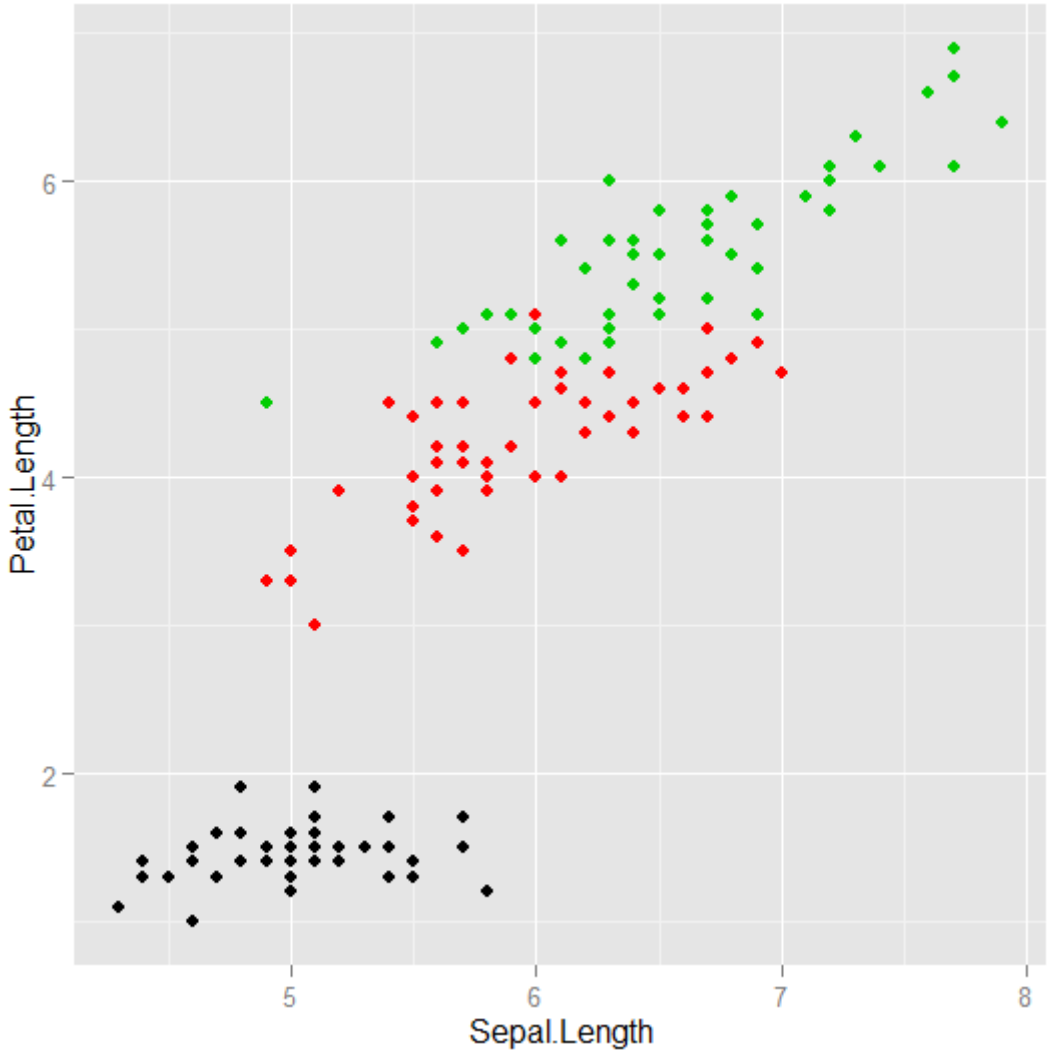
Suppose we find that salary increases with years of education and that this is statistically significant. What confounding variables might there be? Suggest some possible instrumental variables and explain how, and why, we could use them to help confirm (or refute) the idea that more education tends to lead to a higher salary.

6. The table gives data from a (fictional) experiment to determine whether seniors at Marlboro score better than first-years on a standardized math test.

Seniors:	32, 33, 75, 41, 58, 55, 76, 65, 39, 57.
First-years:	46, 35, 36, 29, 48.

What do you conclude from this data?

1. On the level of the entire data set, sepal length is almost certainly correlated with petal length; all three built-in tests of correlation (Pearson's, Kendall's and Spearman's) showed a p value of less than $2.2 * 10^{-16}$. Pearson's assumes a normal distribution, but the other two tests only require the data to be ordinal (which they are), so the result is robust.



In the graph above, this correlation seems to hold true within two of the species (versicolor and virginica), but perhaps not the other (setosa). It's clear that the strength of the overall correlation between petal length and sepal length is a result of difference between species, since species with higher sepal lengths in general have higher petal lengths also.

```
cor.test(iris$Sepal.Length,iris$Petal.Length, method=c("pearson"))

cor.test(iris$Sepal.Length,iris$Petal.Length, method=c("kendall"))

cor.test(iris$Sepal.Length,iris$Petal.Length, method=c("spearman"))


ggplot(iris, aes(Sepal.Length,Petal.Length)) + geom_point(color=iris$Species)
```

2. Three major difficulties in this study are getting a representative control group, making sure people are honest, and nonresponse.

Assumption: I'm comparing Marlboro graduate to graduates of other colleges and universities, not to non-graduates.

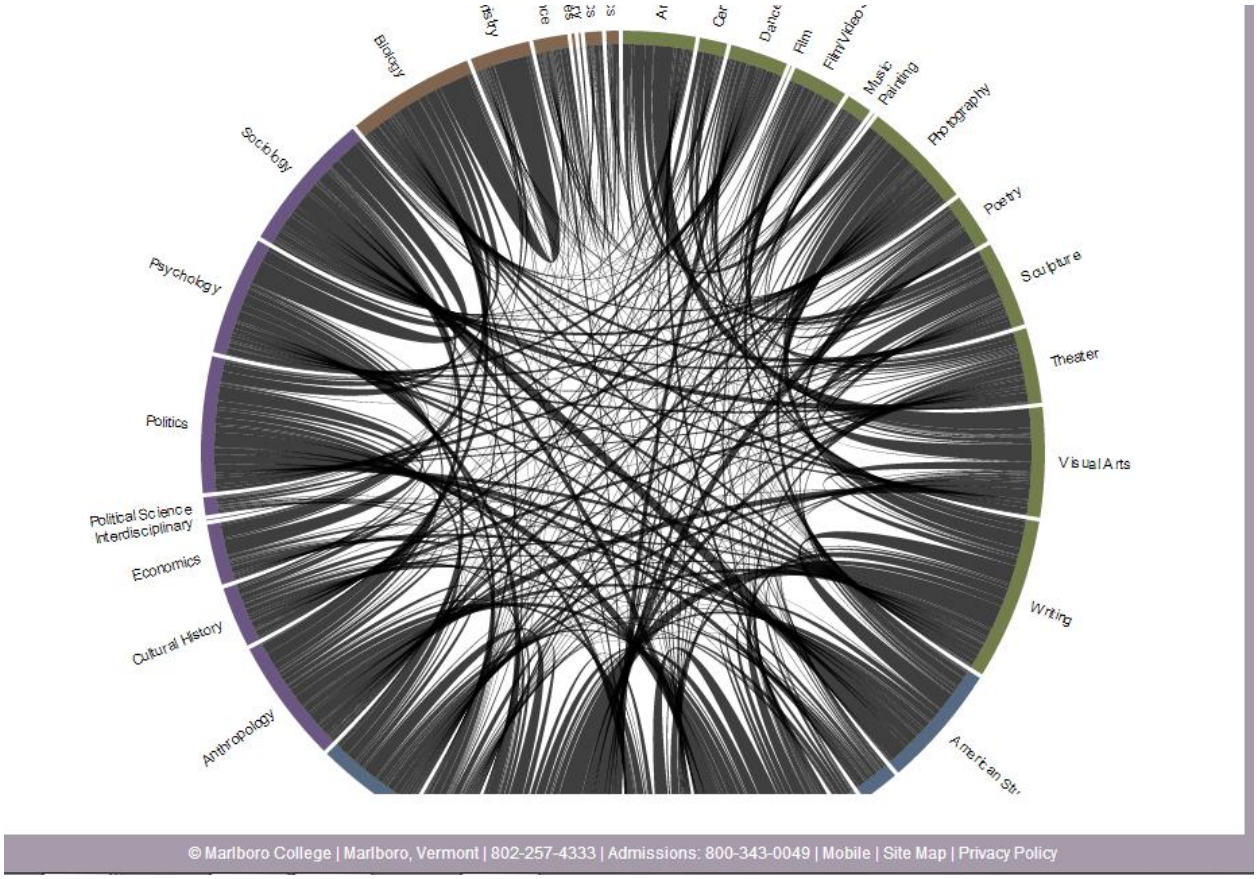
Representative control group: First of all, I want to compare Marlboro students to students from other similar populations, so I need to find figures about our breakdown of gender, race, class, geographical origins, etc. Thus, rich white women from Connecticut (for example) from the Marlboro sample will be compared with other rich white women from Connecticut from other colleges (ideally; due to sample size I might not be able to compare along so many variables at a time). I'd also want the people from outside of Marlboro to be representative of college graduates generally. If possible I would look at nation-wide statistics (e.g. from the Department of Health), assuming they were available and their methodology seemed sound. Otherwise I'd have to survey graduates of a wide variety of colleges myself, which would be cost- and time- intensive.

Honesty: Because drug use is often stigmatized and, in many cases, illegal, Marlboro graduates may be hesitant to admit to drug use (particularly if they don't know me or have reasons to doubt that the data will be confidential). I can deal with this by the instructions in the survey I send them. Essentially, I ask them to flip a coin when answering each question, answering 'yes' if the answer is heads and answering truthfully if the answer is tails. Since I won't know what the coin flip was, they have no reason to be worried about honestly answering 'yes', and so presumably will be honest. One assumption this makes is that they would feel comfortable truthfully answering 'no', which could plausibly be untrue if drug use were considered "cool" or something. Also, this method increases my standard error, so I'd need a larger sample.

Non-response: This is always a problem with surveys, although it only affects the data if people who don't respond are systematically different from those who do. Thus, I have two goals: reduce the non-response rate and try to prevent nonresponse from being related to the answers they would give. For the first part I think it would help to send this survey through alumni director Kathy Waters (whom I have a good relationship with) because she's directly in contact with the most Marlboro graduates. I would also create a lottery for a gift card or Marlboro College-related product (e.g. socks, tote bag), and enter everyone who completed the survey into it, to encourage participation. It's more difficult to deal with the risk of non-respondents being systematically different from respondents, mainly because I don't necessarily know ahead of time what factors would cause that to be true, but I would do my best to make the survey seem non-political so as to avoid alienating any particular group.

3. First graphic: http://www.marlboro.edu/test/plan/16/

On my browser, the bottom part of this is cut off, meaning I can't see anything for most of the humanities. I can get around this by zooming out, although that makes it harder to read and, in any case, a good infographic should be as easily accessible as possible.



The interactivity of it is helpful; when I choose an area of interest (that isn't in the humanities), like Psychology, I can see which disciplines have been in Plans with psychology and which haven't (for example, no one has done a Plan in Psychology and Math), as well as which areas are most commonly linked (unsurprisingly, Psychology is liked to Sociology more often than to any other discipline here). It's also helpful that this links directly to the page of the different disciplines so they can learn more about them, although I think it would be even more helpful if it could link directly to the Plan library also so people could see what a Plan in Psychology and Poetry looks like. I think the organization of disciplines by area is helpful because it makes it easier to use quickly, and the sizing of the different disciplines by number of Plans gives a good idea of what disciplines are the most popular in general. One of the disadvantages of this graphic is that, without looking at a specific discipline, the data look horribly confused and discerning patterns is difficult. While this is information in itself (that Marlboro Plans are very cross-disciplinary, and even cross-area), it would be helpful if something could be done so people could glean more information from the graphic as a whole. I'm not sure what that would be (but feel using color somehow could be helpful), but I have a different suggestion that accomplishes similar goals:

allow users to look at more than one discipline at once. This would allow them to create their own summary information and easily see things like which disciplines often go with Psychology versus with Sociology, or whether Plans in social science disciplines in general are more likely to also be in arts, humanities, or natural sciences. If I had the programming skills to do this, I would definitely allow users to look at multiple disciplines simultaneously (without looking at the entire graphic, as above).

As a final note, I think the width of different disciplines can be slightly misleading, because it makes it seem like very few people do Plans in film, painting, and interdisciplinary studies (when my suspicion is just that very few people in those fields do multidisciplinary Plans). Since I think this width information is useful, perhaps just adding a note somewhere saying that width only reflects volume only reflects Plans that are in more than one discipline would be useful (or having an option of seeing width by total number of Plan instead).

Second graphic (Link added later: http://viz.healthmetricsandevaluation.org/tobacco/):

I should note at the beginning that I was browsing through data visualization blogs last night and skimmed a post about this exact graphic, so my ideas aren't entirely original. The author (I think from FlowingData) said (and I agree) that the smoking prevalence line graph is almost impossible to read. While the average line is helpful and there is limited interactivity, it would be much better if there were some other representative lines, representing either countries of interest or different regions. It's helpful that clicking a country on the map above highlights its trend line, but without selecting specific countries the graphic at the bottom conveys very little information.

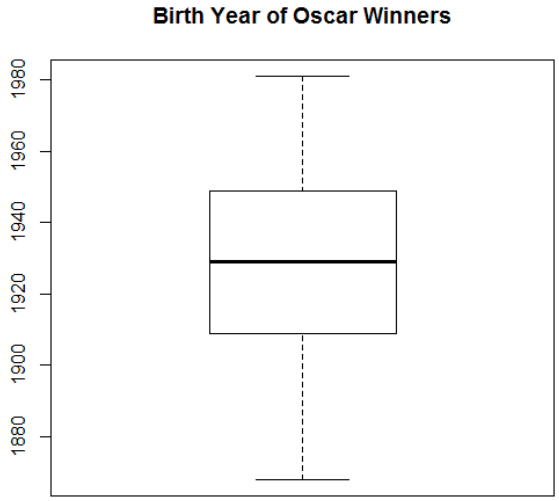
The map at the top is helpful for seeing general geographic trends, although it would be even more helpful if it were a single shade. (Right now it's hard to compare the difference between India and China to the difference between China and Russia, for example.) It might also help to have different blocks (e.g. 0-10%) that were discreet colors; it would definitely be worth trying both ways.

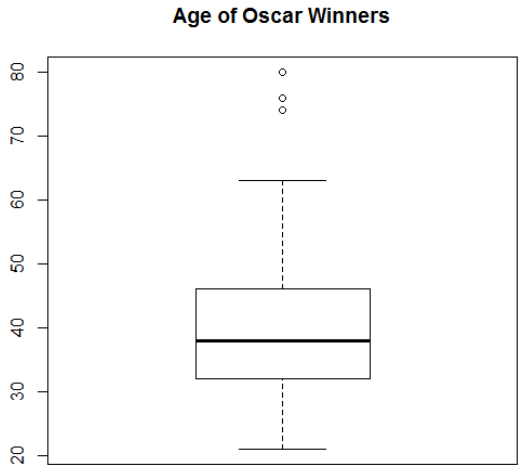
I just ran out of time, so no more comments on this graphic!

4. data=oscars (not oscar)

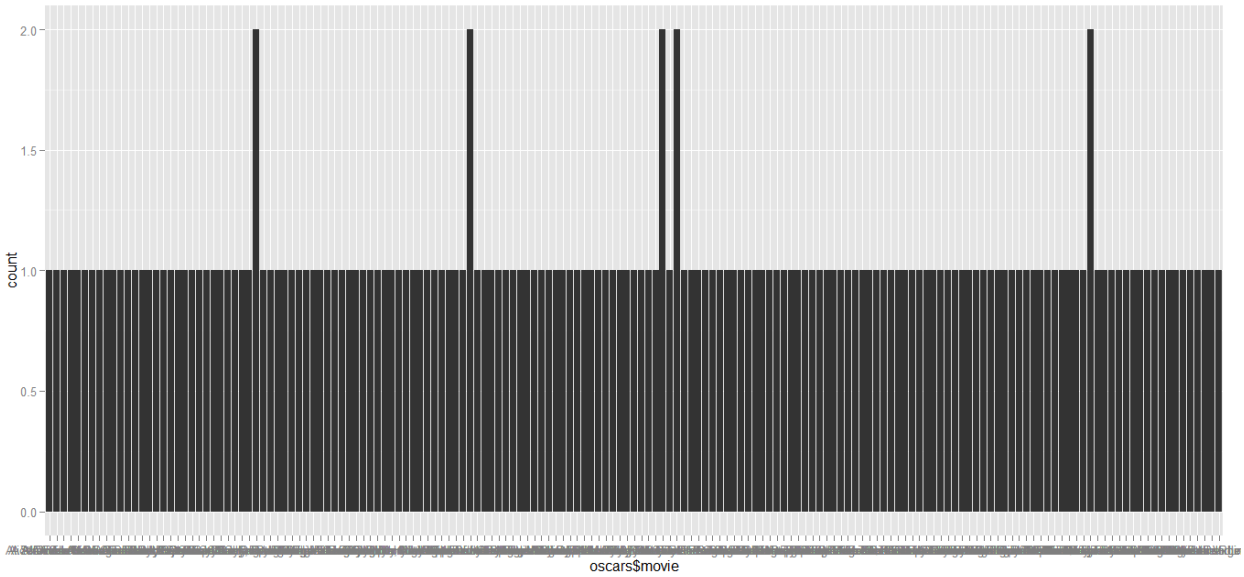
There have been 170 Oscars awards from 1929-2012, with one per sex per year except for two years with ties (females in 1969, males in 1933). Katherine Hepburn is the only person to win four Oscars, while 20 other people have each won the Oscar twice, and no one else has won the Oscar more than once.

The birth years, months, and days are mostly distributed as expected, except that one of the actors' birth months is listed as 26. Since the birth day is listed as seven these could be switched (i.e. it should be July 26th) or it could be a different error, but there was almost certainly an entry error here that really distorts the graph.





According to the really horrible histogram below, there have been 5 movies from which two people won Oscars. (The histogram is impossible to read, but it does quickly show that data and is therefore useful to me as an investigator, even though I'd never show it publicly.)



Finally, it's worth noting that 25 winners came from England (or rather, 25 wins were from English people; an English person who won twice is double-counted here, and similarly for my later statistics), 23 from New York, 22 from California, and 5 or fewer from each other category (states or country). Of

the two with five winners, it worth noting that 4 or Connecticut's five wins come from Katharine Hepburn, while Pennsylvania's 5 wins is actually lower than it should be; the category "Philly" presumably refers to Philadelphia, which means the one win from Philly should be added to Pennsylvania's count.

```
summary(oscars)

summary(oscars$name)

boxplot(oscars$birth_y)+title("Birth Year of Oscar Winners")

boxplot(oscars$birth_mo)+title("Birth Month of Oscars Winners")

boxplot(oscars$age)+title("Age of Oscar Winners")

qplot(oscars$movie)

summary(oscars$birth_pl)
```

5. Some of the major confounding variables are parents' income and their levels of education (these might be separate, as in father_income and mother_income, father_education and mother_education). Plausibly something like state would also be a relevant variable (if some states are rich and focus on education), and it's even possible that something like expected income is a relevant variable (because people who expect to be able to pay off their debts may be more willing to go through college/graduate school). There are countless other possibilities as well (which may not be significant, but we cannot discard a priori) such as malnutrition or good school counselors, that could directly affect both (in the same direction). Hence the need for IV analysis.

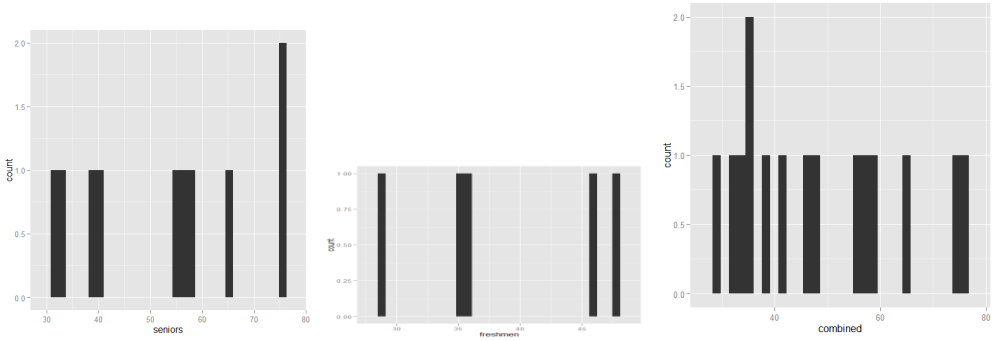
One in this scenario would be looking at changes in school aid. If we look at a major change in state or federal financial aid policy (or multiple policies and a meta-analysis to eliminate random year-to-year changes) and compare students who entered college right before the policy to students who entered college right afterwards[1], we could assume that a difference in aid would affect their level of education, although it's also possible that the lower amount of debt they had would influence their wage.

Alternatively, if historical data is okay, we could look nationwide at when public schools or mandatory education laws were introduced to different cities/states and how that affected their wages. This has been used for crime data (to see how mandatory school attendance affects crime) and is assumed to be exogenous. If introduction of public schools is in fact exogenous, then any difference in wages would be a result of a difference in education, and therefore we could conclude that education is causal. I think

[1] Actually, because this policy might affect students' choice or ability to go to college, it would be better to look at the cohort of all people born 18 years before the policy went into effect, and compare them with the people born 17 years earlier. Then the difference between the two groups could (possibly) be assumed to be due to the policy, and therefore any wage difference would be a result of education.

mandatory school laws are a better IV, although as researchers we may be interested in how schooling affects wages now rather than in the past, and so we'd need to use the more recent data associated with the IV in the paragraph above.

6. With a one-tailed student's t-test (one-tailed because we're testing specifically whether seniors score higher, not whether it's different), we have a p value of 0.01994, so the test is significant at $p<0.05$ but not $p<0.01$. However, this assumes that the distribution of scores is normal. While there are few data points, nothing about any of the three graphs (seniors, freshmen, or combined) looks especially normal; they almost seem closer to a *uniform* distribution.



Seniors *Freshmen* *Combined*

Because of this, and because nonparametric tests work regardless of whether a distribution is normal, the Wilcoxon two-sample test is used instead. Here we get a p value of 0.0646, so $p>0.05$ but $p<0.1$. We therefore cannot conclude at the conventional level of significance that seniors score higher without assuming normality.

```
seniors <- (32, 33, 75, 41, 58, 55, 76, 65, 39, 57)

freshmen <- c(46, 35, 36, 29, 48)

t.test(seniors,freshmen,alternative = c("greater"))

combined <- c(32, 33, 75, 41, 58, 55, 76, 65, 39, 57, 46, 35, 36, 29, 48)
```

```
qplot(seniors)

qplot(freshmen)

qplot(combined)

wilcox.test(seniors,freshmen,alternative = c("greater"))
```

Statistics and Data Visualization: Final Exam 2

Answer all questions. You may use statistical software, a calculator, your notes, books and other sources, but be sure to carefully cite all references. You have one week to complete the exam.

1. Revisit Final Exam 1. With the benefit of more time, is there anything on which you would like to elaborate? Any errors you want to fix? Places where you wish you'd taken a different approach? Do so.

2. There are two data sets called email and email50 that are part of the `openintro` package (see introduction to Final Exam 1 for details of how to get this data into `R`). The data set email contains data on 3921 emails sent to an account, including an indication of whether they were spam. The data set email is a randomly selected 50 entries from the email data set.

Use email50 to build a model for predicting whether an email is spam depending on (some of) the attributes you have available in the data set. Discuss the choices you made and how you came to make them. Test your model on the full data set. How well did your model do? [The skills you've learned in multiple and logistic regression are an adequate base for answering this question, but if you want to explore more and take it further you might try investigating "machine learning".]

3. Along with this exam, you have received three Excel spreadsheets. The first shows how many credits each (anonymous) Marlboro student took in each of the four areas in each semester they were here, along with information about their GPA and plan field. The second two give data on electricity and heating oil use by building for the campus.

Choose one of these two options and create one or more visualizations that help either the Curriculum Committee (in the first instance) or Plant Ops (in the second) understand some aspect of life at Marlboro of interest to them. Also discuss your process for creating the visualization. If you choose the first option, also comment on issues of privacy and confidentiality when dealing with such information.

4. Newspapers are increasingly using data visualization as a tool for conveying aspects of stories to their readers, especially in online versions. Using at least three examples, discuss both good and less good instances of such visualizations and the advantages and disadvantages that come with this trend.

5. You are appointed to the role of Statistics Liaison between the social sciences faculty and the mathematics faculty at a fictional small liberal arts college. Congratulations!

Prepare a discussion piece for both faculties on the topic of parametric vs. non-parametric statistics that lays out the differences between these approaches and your thoughts on how and when these should be addressed within the statistics curriculum to best serve the needs of students in the social sciences. Use of examples encouraged.

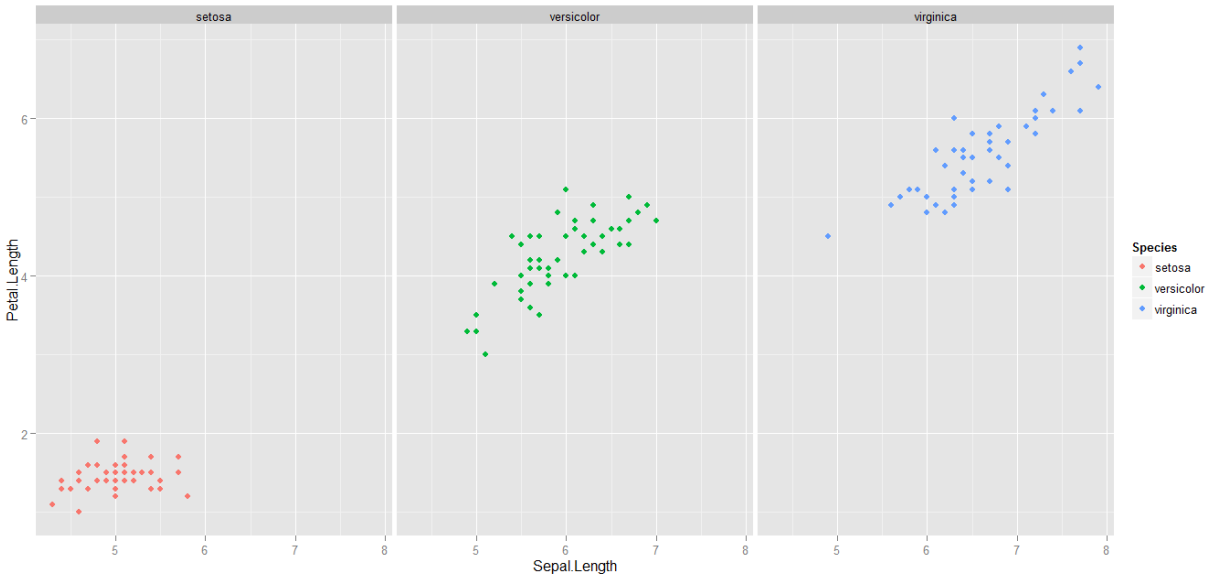
6. Is there a topic on which you were hoping for a question but one did not appear? Or a topic on which you wish you'd had more chance to demonstrate what you can do? Set yourself a question and answer it. You can take it from a textbook, or elsewhere, or make it up from scratch. It can be technical or discursive. It should, however, be a new question for you and not something that you've answered prior to this exam.

5.1 Revisions of Previous Exam

Revisit Final Exam 1. With the benefit of more time, is there anything on which you would like to elaborate? Any errors you want to fix? Places where you wish you'd taken a different approach? Do so.

5.1.1 Flowers

I divided this data set into three parts and tested the correlation in each species using each of the three tests. Setosa showed a weak correlation (around 0.267 for Pearson's), with the p value hovering around 0.05 (a bit higher for Pearson's, a bit lower for Kendall's and Spearman's tests). The other two species had significant correlations (roughly 0.754 and 0.864 for Pearson's), with p values consistently less than 10^{-7} and going as low as $6.661 * 10^{-16}$. Thus, we have weak evidence of a positive correlation between petal and sepal length for setosa and very strong evidence of a positive correlation in versicolor and virginica. A faceted graph below shows each species plotted individually and illustrates both intra-species and inter-species correlations between petal and sepal length.



After trying the help system, I consulted an old file of R code I wrote to see how to facet.

```
setosa <- iris[1:50,]
versicolor <- iris[51:100,]
virginica <- iris[101:150,]


cor.test(setosa$Sepal.Length,setosa$Petal.Length, method=c("pearson"))
cor.test(setosa$Sepal.Length,setosa$Petal.Length, method=c("kendall"))
cor.test(setosa$Sepal.Length,setosa$Petal.Length, method=c("spearman"))
```

```
cor.test(versicolor $Sepal.Length,versicolor $Petal.Length, method=c("pearson"))
cor.test(versicolor $Sepal.Length,versicolor $Petal.Length, method=c("kendall"))
cor.test(versicolor $Sepal.Length,versicolor $Petal.Length, method=c("spearman"))

cor.test(virginica $Sepal.Length,virginica $Petal.Length, method=c("pearson"))
cor.test(virginica $Sepal.Length,virginica $Petal.Length, method=c("kendall"))
cor.test(virginica $Sepal.Length,virginica $Petal.Length, method=c("spearman"))

ggplot(iris, aes(Sepal.Length,Petal.Length,color=Species)) + facet_grid(. ~ Species) + geom_po
```

5.1.2 Data Specification

I talked with Marlboro alumni director Kathy Waters about this; she recommended having this survey go through an outside group to make it more anonymous. I agree, but even with the extra level of anonymity resulting from using this outside group I would probably use the coin-flipping design because people are often hesitant to admit to stigmatized behaviors even on (nominally) anonymous surveys. Both the outside group and larger sample size required would increase costs, but I think the choice should be between not doing the survey and doing it well (but expensively), rather than doing it cheaply or doing it well. (However, if the survey will involve a large number of related questions on the survey there wouldn't be much point to the coin procedure.)

A comment about the original writing. While I didn't note it on the short exam I was intending for it to be anonymous; I don't think it would be especially worthwhile to do the survey otherwise. In my original conception the data would be anonymous but who completed the survey would somehow be logged (separately); Marlboro has done similar things surveying students. In retrospect I don't think this makes sense; it seems better to keep it fully anonymous (and probably not necessary to incentivize alums with Marlboro memorabilia). Anonymity is essential for this study to work well, and for ethical research practices.

Kathy mentioned National Association of Drug Court Professionals (NADCP) as a possible group for data about drug use natiowide; I would investigate them and the Department of Health to see what relevant data they have, and stratify my sample appropriately.

5.1.3 Infographics

I'm mostly happy with my writing about the first infographic, but want to add two things. First, I noticed from looking at it again that a text box shows the exact number of Plans in each area (that have included at least one other area) when the cursor is held over a discipline. This is helpful because sometimes people want to know exact values; it therefore complements the visual approach people already have. (Unfortunately, there seems to be no similar feature to see exactly how many Plan were in two given disciplines, e.g. the number of Plans that combined Psychology and Sociology.) One small negative feature I noticed is that "Interdisciplinary" has a dead link. While this major has less consistent structure, I think it would still make sense for it to go to the Liberal Studies pages. That's definitely a minor thing, though.

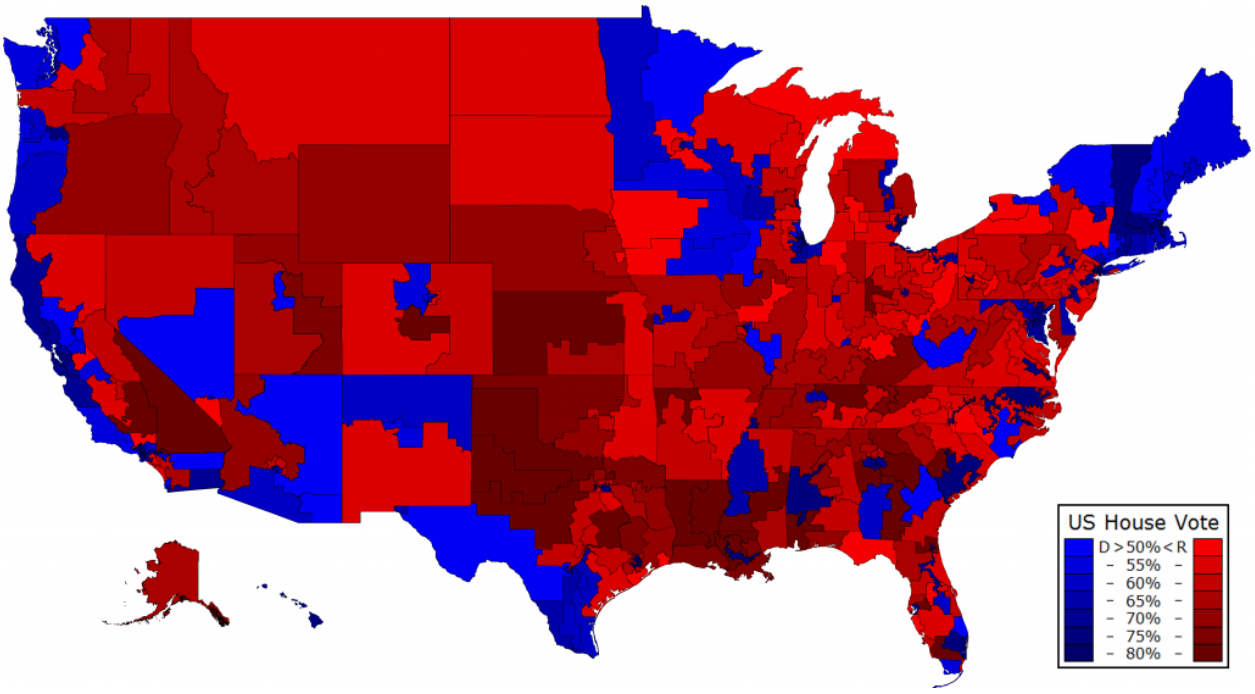
I have much more to say about the second infographic, because I ran out of time in the first exam. Note: the article I referenced in the first exam is actually from junkcharts; it was written by Kaiser Fung:

`http://junkcharts.typepad.com/junk_charts/bumps_chart/`

After having a bit more time to explore this graphic I have much more appreciation for (and awareness of) its interactive features. Being able to filter by age, sex, year, region, and country, and having the option to look at different ages or to see the slideshow over time, adds a lot to the strength of this infographic. (Admittedly, I think I have a definite bias to more features; I support these because they convey information that I think would be of general interest, or at least is to me.) I want to address the rest of the junkcharts article, but first clarify something I said in the original exam.

I still think the map should be monochromatic with shading, so that comparisons between countries currently shown as different hues can occur more easily. The two-color model to me only makes sense if there's a natural middle point that divides the data into two qualitatively different groups (other than the mean/median); that isn't the case here. (As an example, consider the map following this paragraph. The point where support of Republicans and Democrats is equal is a natural and important point, with the two colors representing qualitatively different situations. I would still *consider* using monochrome here (and try it both ways to see which looked better), but this is an example where it is effective with two colors.) Also, by convention, colors should normally get stronger as smoking prevalence (or whatever attribute is being measured) increases; for countries that are blue the opposite is the case, which makes it appear that the U.S. has somewhat *less* smoking than Mexico, and much *less* than Ethiopia. While most readers will check the legend, this makes it harder for them to process unconsciously/preattentively in the way they should. Additionally (because I feel quite strongly about map colors, apparently), I suggest the designer consider using discrete rather than continuous values. From personal experience I've also found that dividing the data into ranges (e.g. 0-10%, 10.1-20%, etc.) and coloring each of them a different shade can be useful because it allows countries in different ranges to be more easily compared[1]; on the other hand, it reduces the amount of information and makes it harder to compare within groups. I could go either way on this one; I think the designer of the graphic should try it if (s)he hasn't.

[1]Part of this has to do with the science of perception. The apparent shade of a color is influenced by the shades of colors immediately surrounding it, so that the apparent smoking prevalence of a country is partially dependant on the prevalence in adjacent countries. Using a small number of discrete color values instead of continuous shading reduces this effect significantly.



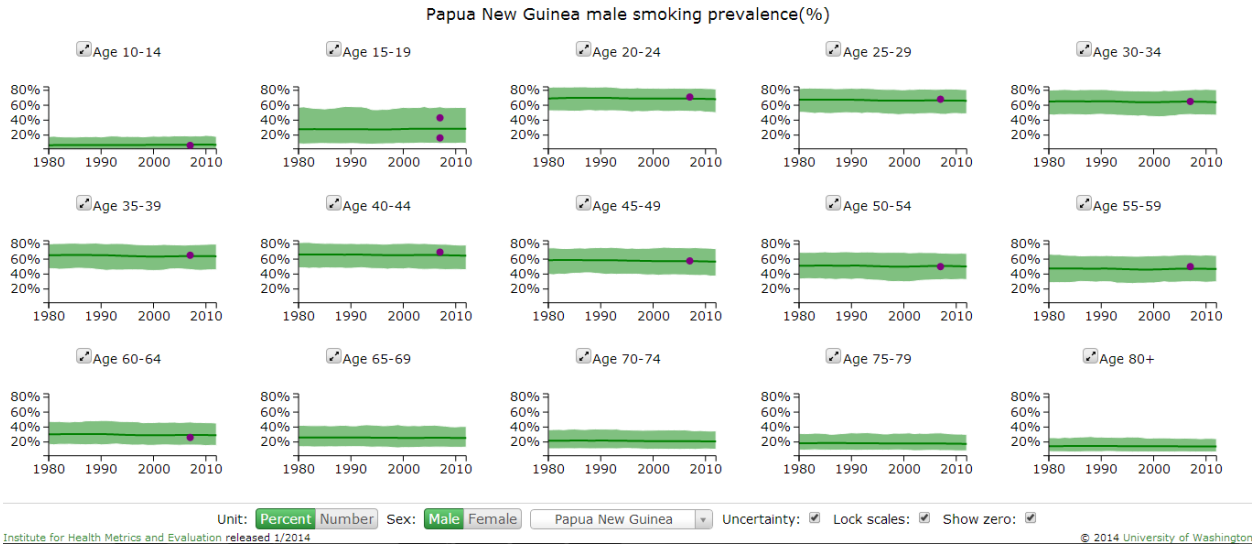
An example map that uses two hues effectively.

`http://www.dailykos.com/story/2013/03/12/1191706/-Political-Geography-Part-5-Mapping-the-US-Congressional-Elections`

I want to respond to two of the points in Fung's junkcharts article. First, regarding the donut chart, Fung talk about how it doesn't make sense for the area to grow larger (because it's on the outside) while the number of people stays the same. (For example, this makes it look like India has more smokers than the whole of South Asia (including India), and the number of smokers in Indonesia looks bigger than Central Europe, Eastern Europe, and Central Asia combined (when it's actually only about 5/9 as large).) Insofar as people are trying to compare across rings this is true, but I think people would mainly be comparing within a given ring, which makes it unimportant. Normally a bar graph would be better for comparing but due to space needs and the inability to easily align all the bars on a single axis while also keeping the nested structure of countries within regions, I think the donut chart works fairly well here. Perhaps a note or some subtle annotation reminding people to not compare across rings would be helpful...

The other criticism Kung has regards the individual data points. This is great for countries like the U.S. and U.K. that have plenty of data, but a lot of countries have minimal amounts of data. (To use an extreme example, Papua New Guinea, pictured below this paragraph, has more trend lines than data points! Despite this, it even allows data to be filtered by sex. (Montenegro is even worse; in Papua New Guinea, *most* of the graphs have at least one data point.)) On the one hand I think these data should still be included; certainly combined data of all the countries in the world gives a better idea of trends than just looking at countries for which a lot of data is available, and aggregate results are pretty reliable even if individual countries' results are not. On the other hand, I don't think there should be trend lines and error bars accompanying these data, because there's really no mathematical basis for that when there are so few data points. Maybe the best bet is to exclude countries without enough data, or maybe instead just the trend lines should be eliminated

(and error bars should only be around the points). I'd probably choose the second option, but either is substantially better than the first. The real problem with this system currently is that it encourages readers to draw conclusions based on little or no data (e.g. the smoking rate of Papua New Guineans ages 25-29 has stayed the same). While there are limited ways to include the data, the way they are included in this infographic is worse than not including them at all. Here I agree with Kung.



5.1.4 Oscars

I was mainly satisfied with this but wanted to find out which films had multiple people win Oscars. I exported it to an excel file and alphabetized the movie names (because I know how to manipulate data better in excel in certain ways); the export code is below.

Films with multiple Oscar wins:

1. As Good As It Gets
2. Coming Home
3. It Happened One Night
4. Network
5. On Golden Pond
6. One Flew Over the Cuckoo's Nest
7. The Silence of the Lambs

Note that the histogram actually missed two of these because of different capitalization made the character strings slightly different (i.e. "As Good As It Gets" and "As Good as It Gets"; "One Flew Over The Cuckoo's Nes" and "One Flew Over the Cuckoo's Nest"). So my original conclusion was wrong: there were actually 7 different films in which two people won Oscars! (Making the histogram was still helpful, though; otherwise I wouldn't have checked for repeat winners.)

```
oscarscsv <- data.frame(oscars)
write.csv(oscarscsv, file = "oscars.csv")

Referenced https://groups.google.com/forum/#!topic/columbiaR/gB2Vih1IW00
for R code about exporting data as a csv file
```

5.1.5 IV

There are many, many IVs for educaton available that one could use nstead of my suggested ones. The instrumental variable I referenced, changes in compulsory education laws (which I actually think works quite well), is from here:

`http://www.clydebankhigh.org.uk/New%20CHS%20Website/Files/modern%20studies/Adv%20Higher/CausesEffects%20of%20Crime/Articles-handouts/education%20and%20crime%202.pdf`

This or any number of other IVs could work, depending on what I had wage data for. The important part is that the variable is exogenous (unrelated to anything in the education model), that it causes a change in education level, and that it only affects wages by affecting education.

5.1.6 Data comparison

Because the standard deviations of the two groups were significantly different (the value for seniors was more than twice as high), I decided to also run the Ansari-Bradley test of variance and the Kolmogorov-Smirnov omnibus test of difference to see if I could at least conclude that the two distribution were different. (Note: because I was running more statistical tests, I would have required a lower p value to accept the conclusion that they were different; otherwise I would artificially increase my likelihood of finding a difference simply by running more tests.) The Ansari-Bradley test had a totally insignificant p value (0.4535!), and the value from the Kolmogorov-Smirnov test was fairly high also (0.1658). In the case of the Kolmogorov-Smirnov test this is most likely the result of a very small sample size (and it being a weak test); the very high value of the Ansari-Bradley variance test could easily be a result of no (or little) difference between variances.

Essentially, my conclusions did not change as a result of these investigations. Still using the standard cutoff for statistical significance, we conclude that the sets are different if we assume the distributions are normal; otherwise we don't have quite enough evidence. Really, though, we should get a larger sample; drawing strong conclusions based on a total of 15 data points in two different samples is only very rarely going to be possible.

```
senior <- c(32,33,75,41,58,55,76,65,39,57)
freshman <- c(46,35,36,29,48)

sd(senior)
sd(freshman)

ansari.test(senior,freshman)
ks.test(senior,freshman)
```

5.2 Modelling Spam

There are two data sets called email and email50 that are part of the openintro package (see introduction to Final Exam 1 for details of how to get this data into R). The data set email contains data on 3921 emails sent to an account, including an indication of whether they were spam. The data set email is a randomly selected 50 entries from the email data set.

Use email50 to build a model for predicting whether an email is spam depending on (some of) the attributes you have available in the data set. Discuss the choices you made and how you came to make them. Test your model on the full data set. How well did your model do? [The skills you've learned in multiple and logistic regression are an adequate base for answering this question, but if you want to explore more and take it further you might try investigating "machine learning".]

5.2.1 email50

After playing around with adding variables to my model I settled on *winneryes*, *numberbig*, *numbersmall*, and *attach*. In general the other variables did not approach any level of statistical or practical significance, with the exception of *format* which was significant until I added *numberbig* and *numbersmall* (they were much better predictors). With those 4 variables I had an R squared value of 0.8074 and an adjusted R squared value of 0.7903. All of the variables (and the intercept) had p values of less than 0.001, except *numberbig*, which was between 0.001 and 0.01.

As a further check, I tried a linear fit model that included every variable measured (except variables like *viagra*, which were the same in all 50 emails). Using that equation, all four of the original variables kept significance values of p less than 0.05, and none of the other variables achieved p values below 0.05 (only one of them, *re_subj*, was below 0.1). Additionally, the original four variable all maintained coefficients with absolute values above 0.25, and only two other variables (*sent_email* and re_subj) had coefficients higher than 0.1 (this is a test of practical, rather than statistical, significance). The R squared value was 0.8468, barely larger than the previous R squared despite the addition of 13 more variables (and a data set with only 50 data points), and the adjusted R squared value of 0.7726 was notably lower. (Adjusted R squared is a measure to account for the fact that the addition of any variable will increase R sqaured; it therefore works to reduce overspecifying.) The original model was therefore better.

I then tried the original four variable model with the addition of *sent_email* and with the addition of both *re_subj* and *sent_email*. The variable *sent_email* is not statistically significant in the first instance, and in the second instance it has a p value between 0.05 and 0.1, while *re_subj* is not statistically significant. The six variable model has an R squared value of 0.8212 (compared with 0.8074) and an adjusted R squared value of 0.7962 (compared with 0.7903), so it is slightly higher on both counts, but not by much. In this model *sent_email* had a coefficient of -0.11486 and *re_subj* had a coefficient of only 0.08901. Since these two variables had relatively small practical and statistical significance, and because the change in the adjusted R squared was small (less than 1%), I decided to omit them from the model.

A few of the variables (*from*, *image*, *inherit*, *viagra*, and *urgent_subj*) had the same value at every point in the 50-email sample. Thus, regressing them for the 50 email sample was impossible, but they may be important variables for the population as a whole (and their value doesn't effect the results in the smaller sample). I would guess that *from* will stay constant and that *viagra* and *urgent_subj* will correlate positively with *spam*, but these are entirely for sociological reasons rather than any reading from the data (I'm not sure about *image* and *inherit*). Since I didn't have any

reason from the data to include them, and since I had no way to guess what their magnitude would be if they *were* included, I decided to not put them into the model. Thus, the linear equation I used is below:

$$spam = 1.00000 winneryes + -0.26508 numberbig + -0.26691 numbersmall + 0.46618 attach + 0.26691$$

A couple final notes: the residual standard error for this equation on the 50 email data set was 0.1368, with the residuals ranging from -0.2669 to 0.7331 The median value, which was also the value at the 1st and 3rd quartiles, was -0.001828; this corresponds to *numbersmall*=1, with for *winneryes*, *attach*, and *spam* all 0. Since both *winnerbig* and *winnersmall* almost exactly cancel out the intercept, and since their values are almost identical, I suspect that their difference in values if due to overspecifying (and that I could model this with an intercept of 0 and a variable called *nonumber* corresponding to having 0 for both *numberbig* and *numbersmall*. However, since the difference between the two variables is minimal (less than 0.4%), I decided to ignore that and leave the model as above. (I should note that, given that this sample has only 5 spam emails and is missing multiple variables, I didn't have especially high expectations for the model. The R squared value will almost certainly be significantly lower, although still positive.)

```
summary(email50)

summary(lm(email50$spam ~ email50$winner + email50$number + email50$attach))

summary(lm(email50$spam ~ email50$winner + email50$number + email50$attach +
email50$to_multiple + email50$cc + email50$sent_email +
email50$time + email50$attach + email50$dollar + email50$password +
email50$num_char + email50$line_breaks + email50$exclaim_subj +
email50$re_subj + email50$format + email50$exclaim_mess))

summary(lm(email50$spam ~ email50$winner + email50$number + email50$attach + email50$re_subj +
email50$sent_email))

summary(lm(email50$spam ~ email50$winner + email50$number + email50$attach + email50$sent_emai

summary(residuals(lm(email50$spam ~ email50$winner + email50$number + email50$attach)))
```

5.2.2 email729

With my model in place I tested it on an intermediate number of emails; I arbitrarily chose the first 729 from the larger set of 3,921 (I like the number 729 because of 1729). Because there are some things I do better in excel I exported the email data to a csv file and then created some new data columns. I created "numberbig" and "numbersmall" so I'd have more control over the number variable in regression, and the variable "winneryes"; all three of these variables were numeric (1 or 0) and allowed me to regress in excel. I then created a column with the formula I had used before to see how well it predicted spam on this larger set; it looked like this:

```
=SUM(Y2,-0.26508*X2,-0.26691*W2,0.46618*H2,0.26691)
```

Finally, I created a residuals column to see how far off my prediction was (and another column that gave the absolute value of the residuals so they could be averaged). I then loaded this new file into R and created the a data set from only the first 729 emails.

In the original data set the residual standard error had been 0.1388, which the residuals ranging from -0.2669 to 0.7331 and the first and third quartiles both -0.001828; the mean (of the absolute values) was 0.08128. Testing the equation from the original model on email729 yielded residuals ranging from -9.324 to 1.000, with the same median and quartile value of -0.00183 and a mean absolute value of 0.09982. For most of the data this was fine, and the mean was relatively unchanged, but the range increased enormously (and the variability somewhat), which is what would be expected from having more data. As a check, the residual standard error of a model using the same variables on email729 was 0.1969, which is about 50% higher. That's quite similar given a 14-fold increase in the amount of data points, although some of the parameters changed quite significantly (in general they were around 35% as large as previously, and attach was no longer statistically or practically significant) so this underpredicts the error rate of my model with the original parameters. My three variables numberbig, numbersmall, and winneryes all remained statistically very significant (and practically noticeable), which was a plus. The R-squared value of the new model, however, was merely 0.08216, about a tenth as large as the value for email50 (despite the fact that the paramters were automatically adjusted); clearly I needed to redo the model with new (now relevant) variables.

Exerimenting as I had to create the first model, I created a new model for email729 that included the variables *winneryes*, *numberbig*, *numbersmall*, *attach*, *from*, *image*, *viagra*, *to_multiple*, *format*, *dollar*, and *re_subj*. About half of the new variables were things that couldn't be included in the previous model because their value had been constant (e.g. nothing in email50 had contained *viagra*), while other variables were added because they had coincidentally not been significant in the first 50 emails but were in the broader category. With the excepion of *dollar* (on both counts), all of them had p values below 0.05 and the coefficients were higher than 0.034. (The p value of *dollar* was a bit less than 0.1 and the coefficient was only about 0.0027, but I decided to keep it due to the relatively large sample size and the fact that the variable dollar wet as high as 54.) The R-Squared value for this model was 0.1746, and the adusted R-Squared was 0.1619. (So it was about twice as high as with only 4 variables, but (predictably) much lower than the R-Squared value for email50.) Here is the model I used:

$$spam = 0.359358 * winneryes - 0.084484 * numberbig - 0.086794 * numbersmall + 0.052865 * attach$$
$$-0.961417 * from - 0.053407 * image + 0.120851 * viagra - 0.067030 * to\_multiple$$
$$-0.034488 * format - 0.002696 * dollar - 0.040945 * re\_subj + 1.121282$$

The residual standard error was 0.1877, with residuals ranging from -0.51922 to 0.99396.

```
emailcsv <- data.frame(email)
write.csv(emailcsv, file = "email.csv")
```

```
emaila <- read.csv("email.csv")

email729 <- emaila[1:729,]

summary(lm(email729$spam ~ email729$winner + email729$numberbig +
email729$numbersmall + email729$attach))

summary(email729$residual1)
summary(email729$residual)

summary(lm(email729$spam ~ email729$winneryes + email729$numberbig +
email729$numbersmall + email729$attach + email729$from +
email729$image + email729$viagra + email729$to_multiple + email729$format +
email729$dollar + email729$re_subj))

summary(lm(email729$spam ~ email729$winneryes + email729$numbersmall +
email729$attach + email729$to_multiple + email729$cc + email729$sent_email +
email729$numberbig + email729$attach + email729$dollar + email729$password +
email729$num_char + email729$line_breaks + email729$from + email729$image +
email729$viagra + email729$re_subj + email729$format + email729$exclaim_mess))

summary(residuals(lm(email729$spam ~ email729$winneryes +
email729$numberbig + email729$numbersmall + email729$attach +
email729$from + email729$image + email729$viagra + email729$to_multiple +
email729$format + email729$dollar + email729$re_subj)))
```

5.2.3 email3921

For my final test I put this equation into the same spreadsheet and created another pair of residuals columns; the linear equation looked like this in excel:

```
=SUM(0.359358*Y2,-0.084484*W2,-0.086794*X2,0.052865*I2,-0.961417*D2,
-0.053407*H2,0.120851*M2,-0.06703*C2,-0.034488*Q2,-0.002696*J2,-0.040945*R2,1.121282)
```

I then uploaded the data again and tested the model from *email729* on the entire data set. (I also ran a check to make sure I had entered everything correctly in excel; good thing too, because the first time I entered something incorrectly!) The range of residuals was significantly narrower han from the original model, going from -1.000989 to 1.031513, athough the average value was only slightly lower: 0.1244 instead of 0.1405. In that sense it worked as a farily good predictor.

I then tried linearly modelling the entire data set from scratch to see whether I had specified the right variables. When I did this, all of the variables were statistically significant with a p value below 0.01 (and most of them had p values less than 10^{-5}). All of the parameters changed, although probably a majority were similar values. (Significant changes don't necessarily indicate weakness in the model; it coud just mean that a different system works slightly better. However, having similar values indicates strength.) The R-Squared value was 0.1452 and the adjusted R-Squared was 0.1428; this is lower than the 0.1746 and 0.1619 values for *email729*, but not by a huge amount.

(The residual standard error was 0.2697.) This indicate my model running relatively well on the large data set, unless another model could do much better.

To conclude, I tried running a model that included every variable possible[2]. With the exception of *dollar* and *re_subj*, which had p values between 0.05 and 0.01, all of my variables had p values less than 0.001. This model included the variables *password*, *exclaim_mess*, and *urgent_subj* as also statistically significant. The R-Squared value of 0.1736 was notably higher (by about 20%), as was the adjusted R-Squared (0.1698, around 17% higher). The residual standard error, however, was extremely similar at 0.2654. Overall I think the predictions were fairly, and that my model performed within the same ballpark as the (overspecified) model with every available variable. This modelling has also shown me that, unfortunately, it's difficult to predict which emails are spam!

```
emaila <- read.csv("email.csv")
summary(emaila)

summary(lm(emaila$spam ~ emaila$winner + emaila$numberbig +
emaila$numbersmall + emaila$attach + emaila$from +
emaila$image + emaila$viagra + emaila$to_multiple + emaila$format +
emaila$dollar + emaila$re_subj))

summary(lm(emaila$spam ~ emaila$winneryes + emaila$numbersmall +
emaila$attach + emaila$to_multiple + emaila$cc + emaila$sent_email +
emaila$numberbig + emaila$attach + emaila$dollar + emaila$password +
emaila$num_char + emaila$line_breaks + emaila$from + emaila$image +
emaila$viagra + emaila$re_subj + emaila$format + emaila$exclaim_mess +
emaila$urgent_subj))
```

[2]I excluded time from this, because for some reason R treated time as a qualitative variable and therefore created a perfect model (because each email had a different time). Since this system has a variable corresponding to each individual data point (and some extra variables), it is a perfect example of overspecifying; we would have no reason to expect it to generalize well. Thus, I included every *other* variable.

5.3 Breadth of Study at Marlboro

Along with this exam, you have received three Excel spreadsheets. The first shows how many credits each (anonymous) Marlboro student took in each of the four areas in each semester they were here, along with information about their GPA and plan field. The second two give data on electricity and heating oil use by building for the campus.

Choose one of these two options and create one or more visualizations that help either the Curriculum Committee (in the first instance) or Plant Ops (in the second) understand some aspect of life at Marlboro of interest to them. Also discuss your process for creating the visualization. If you choose the first option, also comment on issues of privacy and confidentiality when dealing with such information.

The dataset I was given for this actually contains quite a bit of information, easily enough that I could probably identify any student I wanted in it despite the lack of names. (It helps that most departments have only a few graduating seniors each semester, and frequently only one!) As is reaosnable the school has a policy about giving these data to students and I had to sign a confidentiality form to access it. Given that process, the more interesting question regards what I can do with them information in terms of sharing it. The Curriculum Committee includes student members and their meetings are open (at least parts of them), so even if I only show this to that Committee I have to make sure that I don't give out any information that could be identifying. Cerainly I can't share the entire data set! This puts a limit on the types of statistical analyses I can do.

Fortunately, the information I found most useful and interesting was not correlational; it was based on analyzing only one variable at a time. Essentially this entire infographic is made up of box plots and violin plots (which are a sort of combination box plot/histogram). For these data, the only way someone could identify a student is by knowing a piece of data about them (e.g. my friend took 131 art credtis; she's at the top of the box plot), but since this isn't linked to any other information, they would be unable to derive any additional information about that student. Thus, while I need to make sure to not share the original data set with other students, there is no apparent risk to sharing the infographic. (Just to be safe, though, I avoided mentioning the years that are being analyzed; I'll tell Curriculum Committee directly if they decide it's relevant.)

The infographic, which follows the next paragraph, makes heavy use of violin plots because I believe they show the data significantly more effectively than alternatives do. This is not an especially commonly-used type of graph, and it's likely some members of Curriculum Committee will be unfamiliar with it. I try to explain how it works in the infographic itself (and present my first set of data as both a violin plot and abox-and-whiskers plot), but there may still be confusion because of unfamiliarity. In this particular case I thought it was worthwhile to do this way anyway, in part because my intended audience is small enough that I can clarify any confusion in person if necessary (although that's definitely not ideal), or they can in the Curriculum Committee meeting. If it were for a larger audience I would probably still use violin plots, but I'm not positive; using a new type of graphic that might not be intuitive is always risky, even if I explain it fairly well.

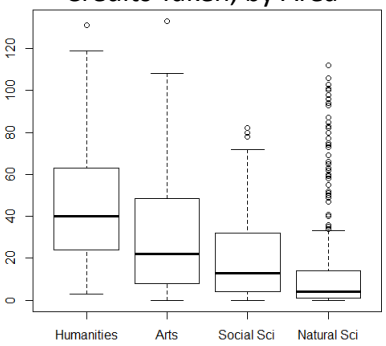
One other small note is that I claim that social science students study more broadly. In the context of the infographic this appears to be a conclusion drawn solely from the graph; in reality I also ran a regression model to see how well the number of social science credits (*ssc_total)* predicted total breadth score (*breadth*). The R-Squared value was approximately 0.098; in other words, roughly 10% of the total variation in breadth score can be predicted from the number of social science credits a student takes. This is related to, but not a logical consequence of, social science

having (by far) the lowest maximum number of credits; no one took more than 82.

Academics at Marlboro

Popularity of Different Areas at Marlboro

Credits Taken, by Area

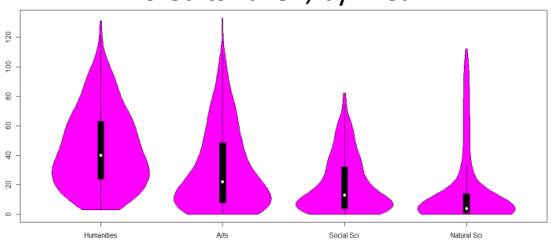


This box plot shows the distribution of credits in each area, from 0 to 133.

The box plot on the left shows that Humanities is markedly the most popular area at Marlboro, followed by Arts, Social Sciences, and Natural Sciences. Some elements of note:

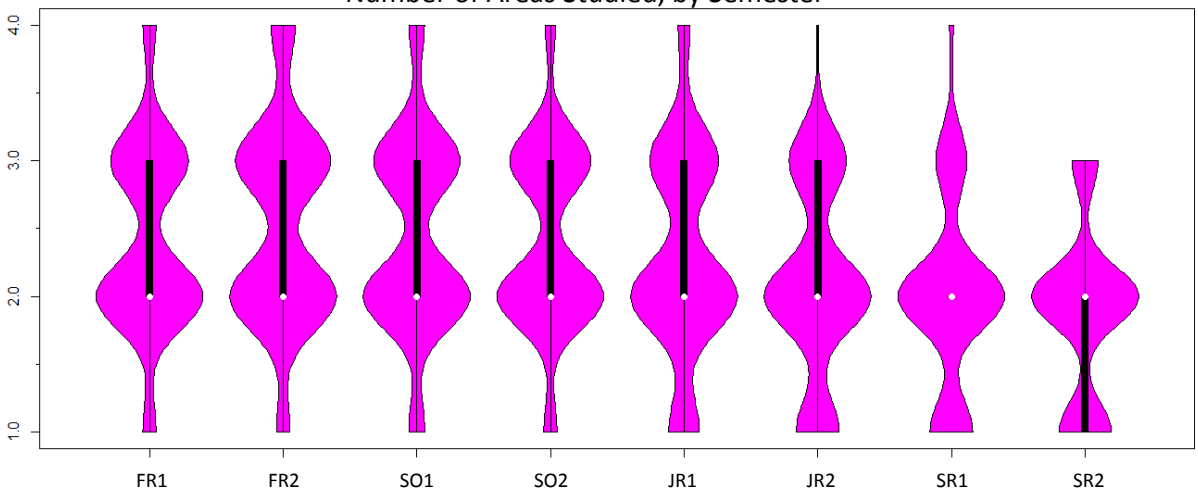
- Every graduate in the data set has taken at least 3 Humanities credits
- At least ¾ of students have taken fewer than 15 credits in Natural Sciences; one result is that almost all NSC majors (with 34+ credits) are outliers
- SSC students study more broadly than students from other areas

Credits Taken, by Area



This violin plot represents the same data as the box plot on the left. The yellow dot is the median, the black bar is the range between the first and third quartile, and the thickness shows frequency for each number of credits. Violin plots, in general, show more information; here it is clearer that the large number of outliers in the natural sciences is due to the unusually low center, a result of having comparatively few NSC majors.

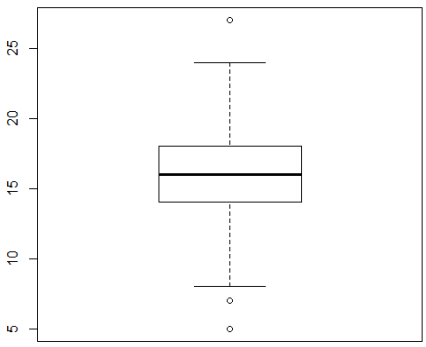
Number of Areas Studied, by Semester



This graph shows the number of different areas people study in during each semester. Every student who takes at least one class in a given semester is included for that semester. A number of interesting trends are visible:

- The median number of areas studied is two every semester, but at least 25% study three or more every semester before senior year
- Some students continue to take classes in each area in their SR1 semester, but it declines rapidly starting JR2 and disappears by SR2
- The number of students taking classes in only one area is small until they start Plan (JR1), and only reaches 25% at SR2 semester

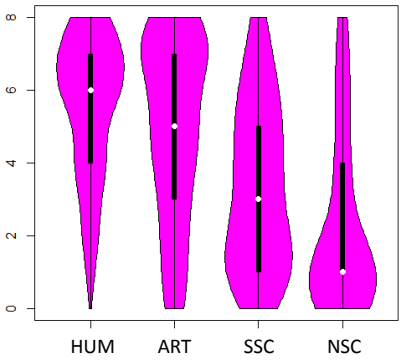
Breadth Score



Because these data include transfer students, some scores are lower than 8; here the data range from 5 to 27.

These two graphs show breadth of study across students' entire time at Marlboro. On the left, the number of areas students take classes in each semester are added together, generating a breadth score from 8-32 (or lower for transfer students). The graph on the right shows the number of semesters a student takes a credit in each area, 0-8. As might be predicted from above, most students (75%) take fewer than five semesters of NSC classes, and a majority do one semester or fewer. On a more positive note, a majority of students spend more than half their semesters taking arts courses; ditto for humanities!

Semesters Studying, by Area



Note: I modified the labels in the original excel file to make it easier for me to use. I also created categories like art_total, which are the sum of the 8 art semesters, and fr1_breadth, which is the number of areas studied in as a Freshman 1. Finally, I created a measure within each area of how many semesters the student took a class in that area, and then a total "breadth score" representing the sum of the individual breadth scores.

```
Referenced http://cs.marlboro.edu/courses/spring2013/painting/closed/
closedresources.attachments/admissions.r to see how to remove data when values were too low.

Referenced http://www.statmethods.net/graphs/boxplot.html for violin plots

Referenced http://stackoverflow.com/questions/2739159/inserting-a-pdf-file-in-latex
to include pdf infographic in LaTeX



study <- read.csv("broadstudy2.csv")

study2 <- study[study$total > 119, ]

art_total <- study$fr1_art + study$fr2_art + study$so1_art + study$so2_art +
study$jr1_art + study$jr2_art + study$sr1_art + study$sr2_art
hum_total <- study$fr1_hum + study$fr2_hum + study$so1_hum + study$so2_hum +
study$jr1_hum + study$jr2_hum + study$sr1_hum + study$sr2_hum
nsc_total <- study$fr1_nsc + study$fr2_nsc + study$so1_nsc + study$so2_nsc +
study$jr1_nsc + study$jr2_nsc + study$sr1_nsc + study$sr2_nsc
ssc_total <- study$fr1_ssc + study$fr2_ssc + study$so1_ssc + study$so2_ssc +
study$jr1_ssc + study$jr2_ssc + study$sr1_ssc + study$sr2_ssc

#boxplot(hum_total, art_total, ssc_total, nsc_total,
#names=c("Humanities","Arts","Social Sci","Natural Sci"))


study729 <- study[study$fr1_breadth > 0, ]
fr1_breadth <- study729$fr1_breadth
study729 <- study[study$fr2_breadth > 0, ]
fr2_breadth <- study729$fr2_breadth
study729 <- study[study$so1_breadth > 0, ]
so1_breadth <- study729$so1_breadth
study729 <- study[study$so2_breadth > 0, ]
so2_breadth <- study729$so2_breadth
study729 <- study[study$jr1_breadth > 0, ]
jr1_breadth <- study729$jr1_breadth
study729 <- study[study$jr2_breadth > 0, ]
```

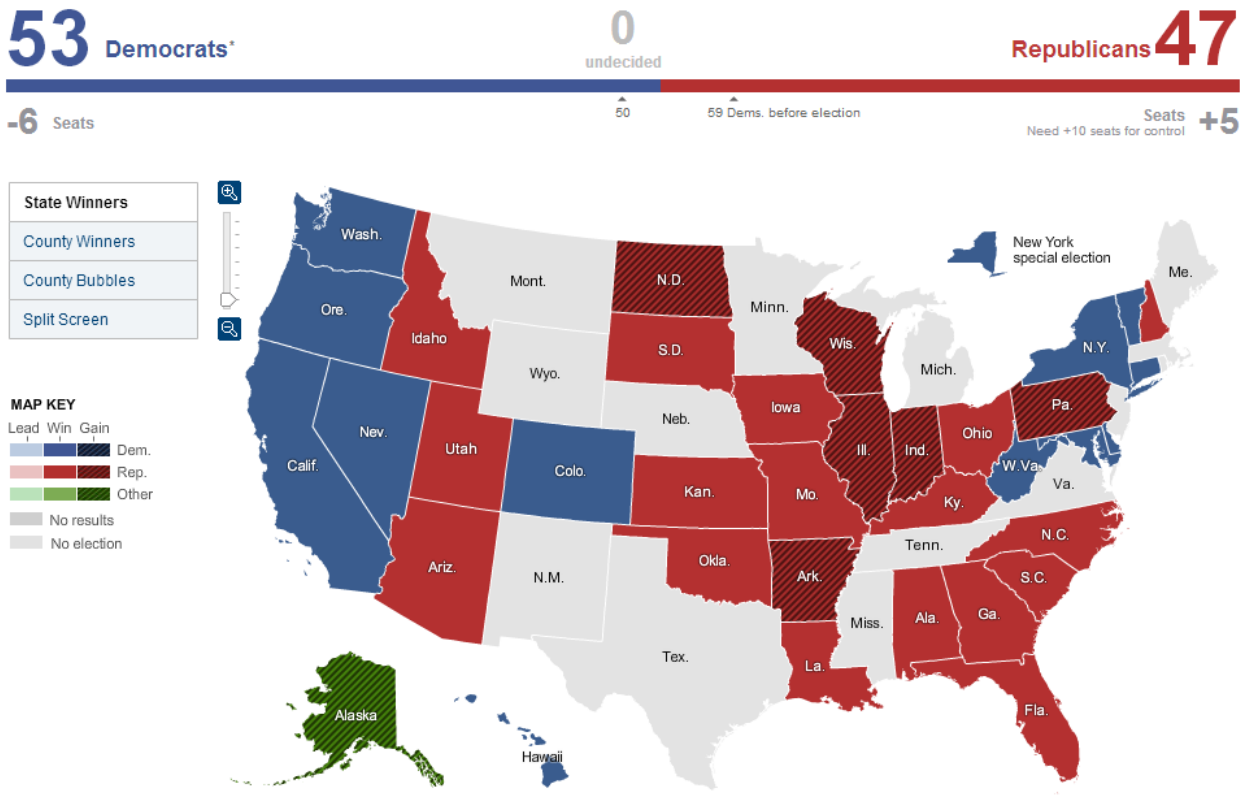
```
jr2_breadth <- study729$jr2_breadth
study729 <- study[study$sr1_breadth > 0, ]
sr1_breadth <- study729$sr1_breadth
study729 <- study[study$sr2_breadth > 0, ]
sr2_breadth <- study729$sr2_breadth



install.packages("vioplot")
library(vioplot)
vioplot(fr1_breadth,fr2_breadth,so1_breadth,so2_breadth,
jr1_breadth,jr2_breadth,sr1_breadth,sr2_breadth)
boxplot(study2$hum_total, study2$art_total, study2$ssc_total, study2$nsc_total,
names=c("Humanities","Arts","Social Sci","Natural Sci"))
vioplot(study$hum_total, study$art_total, study$ssc_total, study$nsc_total,
names=c("Humanities","Arts","Social Sci","Natural Sci"))

study <- study[study$breadth > 0, ]
cor.test(study2$ssc_total,study2$breadth, method=c("pearson"))
cor.test(study2$total,study2$breadth, method=c("pearson"))
boxplot(study2$breadth)

vioplot(study$hum_breadth,study$art_breadth,study$ssc_breadth,study$nsc_breadth)
```

5.4 Data Visualization in Newspapers

Newspapers are increasingly using data visualization as a tool for conveying aspects of stories to their readers, especially in online versions. Using at least three examples, discuss both good and less good instances of such visualizations and the advantages and disadvantages that come with this trend.

In general the main advantages of this trend are the ability for readers to see more information (more quickly, easily, etc.)and creating more interest; the disadvantages are that badly made data visualizations can obscure information and the trend may draw interest from word-heavy articles. I'll present some examples first and then discuss these advantages and disadvantages in more detail. (As a note, most of my screenshots show only part of the infographic in question.)

First, consider this infographic from the NY Times:

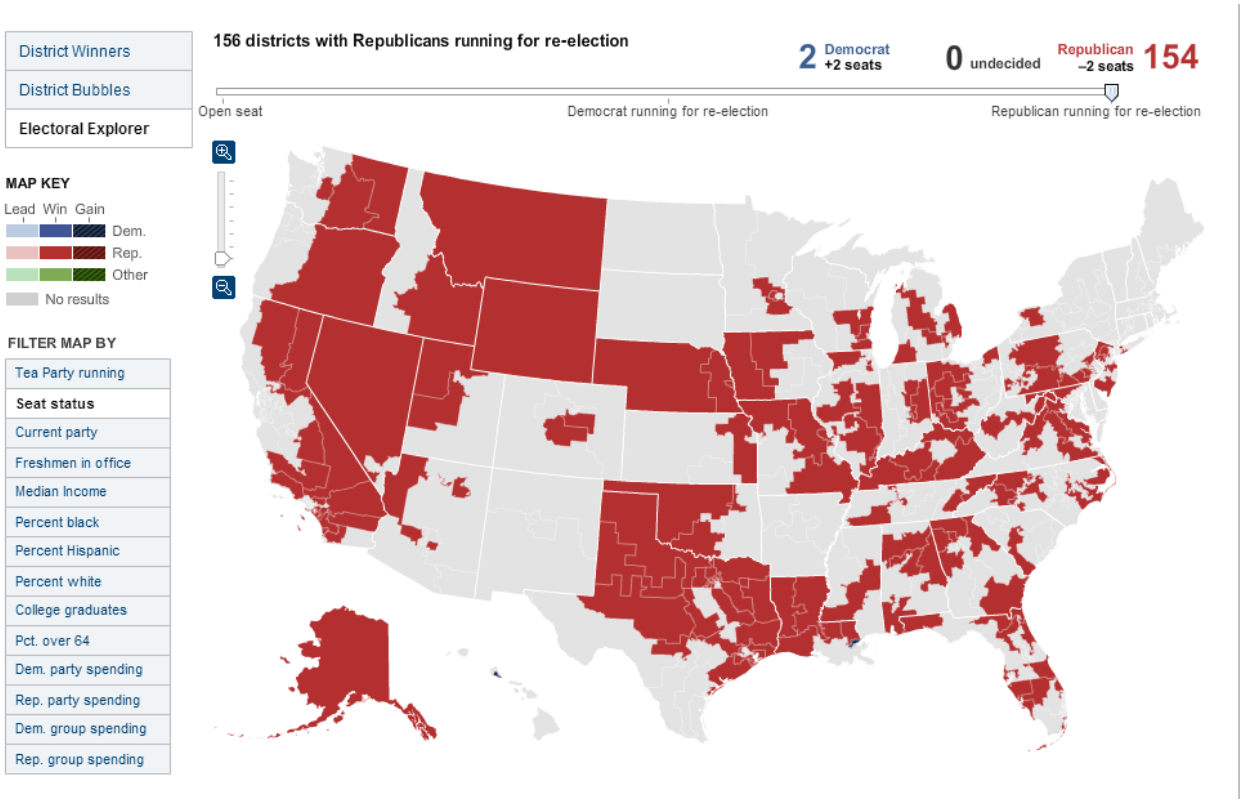


`http://elections.nytimes.com/2010/results/senate`

This infographic shows the Senate races of 2010. The key marks which party won, and whether a different party had had that seat before the election (while votes were being counted, they also showed which party was winning). The data can be examined at state level or county level, or can be compared with historical elections from 2002-2008. Holding the cursor over a state causes a box to pop up showing the exact number of votes for each candidate. This picture shows senate races but the map is interactive and also allows users to see house races or governor races, or pick a state and see all of those races at once. Users can also look at a table version of the data, see what happened hour-by-hour, and look at the demographic breakdown of voters from exit polls.

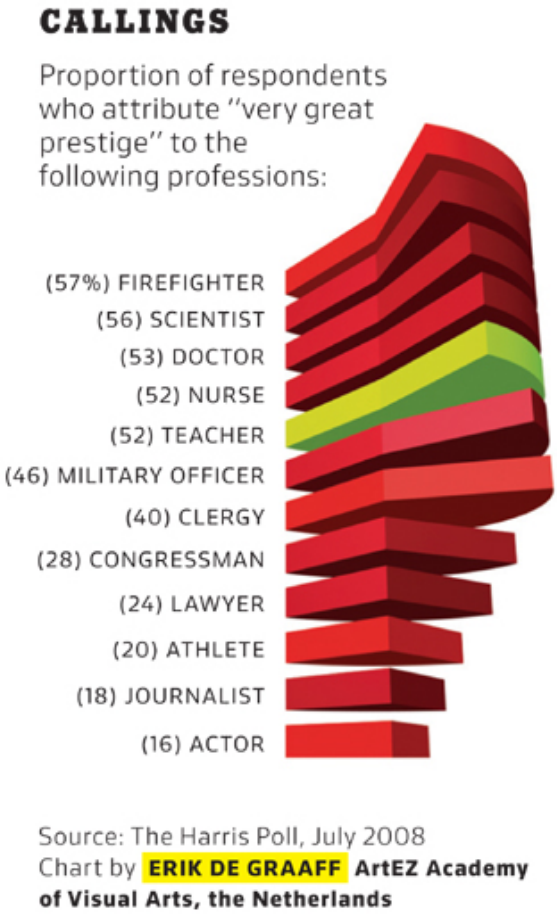
In short, there are many different options for users, and the options are both clearly organized and interesting to many people.

I think this infographic is exceptionally good. It's quite engaging and usable, with a lot of interesting and varied information, and so it encourages users to explore whatever parts of the election are of interest to them. No sort of text representation would be nearly as effective (the tables also use principles of infographic design, and colors) or concise at conveying information. The interactive elements are very intuitive, and this visualization was constantly being updated as votes were being counted, adding a certain level of drama to it that was exciting! One thing I discovered was just how dominant Republicans were in the House; while Republicans won a fair number of seats against Democratic incumbents, virtually no Democrat was able to defeat a Republican incumbent (the two counties where it happened are so small they're not immediately visible on the map below). This map uses the electoral explorer feature, an addition for house races that allow users to explore the data with even more depth.



The two districts where the Democrat defeated a Republican incumbent are in Hawaii and Louisiana on this map

The NY Times is also capable of some really poor infographics, as below:



`http://www.nytimes.com/2008/09/21/magazine/21wwln-lede-t.html?pagewanted=all&_r=0`

This graph is a combination line graph and pie chart, in three dimensions. Comparing areas in pie charts is notoriously difficult (or ineffective); doing so in a normal 3-D pie chart is worse. In order to fit the pie slices into something resembling a line graph, the author here has chosen one of the worst possible angles to view a 3-D pie chart, making it even more difficult to read. The upshot is that comparing volumes precisely is quite challenging, and doing so preattentively is likely to give the reader an inaccurate impression of the relative percentages for each profession. (Indeed, I would have no idea of their values were it not for the numbers on the left that completely ruin the poit of making an infographic.) While I think this visualization is substantially worse than a table, it is worth noting three things the designer did correctly: the labels are horizontal rather than vertical, the values are ordered, and the profession of interest (teaching) is marked in a different color so as to stand out. Overall, though, this visualization distracts from both the text of the article generally and the information it (the visualization) is trying to convey.

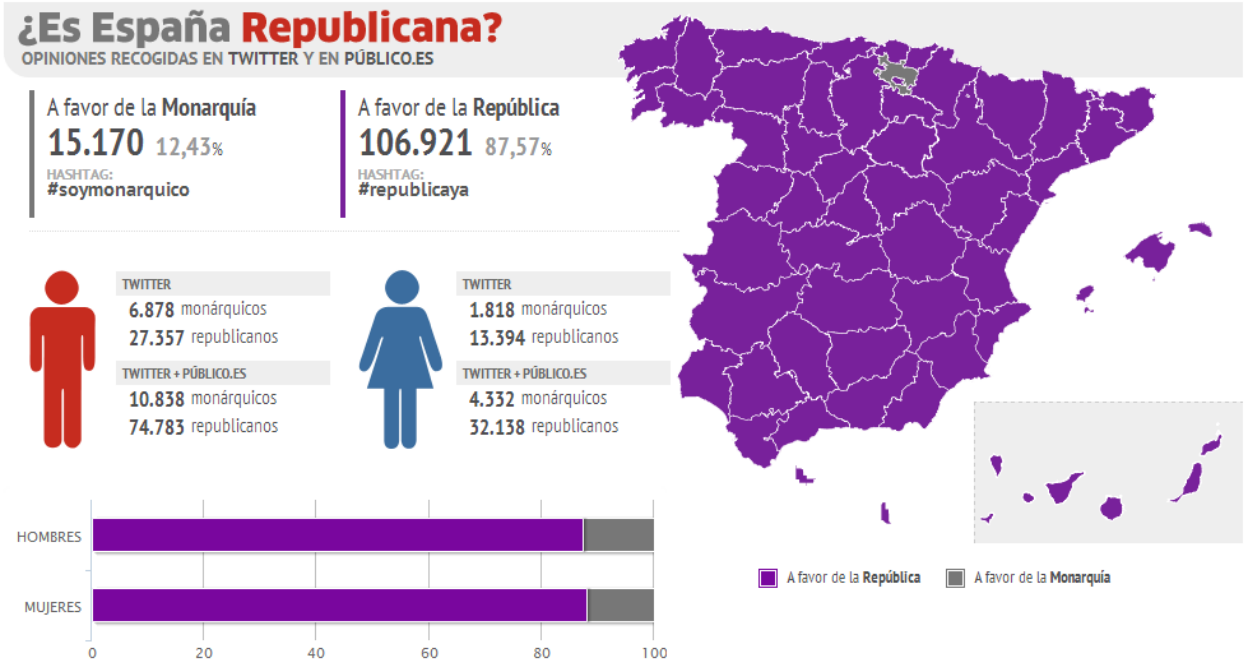
Overall I think the trend toward including more data visualizations in (respectable) newspapers is a positive one. As in the NY Times example above, it conveys much more information than text,

organizes and filter it, responds to individual interests, and engages readers. I can appreciate it as someone who is *not* an especially visual learner; for people who think more visually, everything I'm claiming is even more true. I also think interest is noteworthy and more important than it might be given credit for; having a really strong, engaging data visualiztion will encourage more people to read news, and to spend more time looking at each article deeply (when done well).

Interactive data visualization in particular are able to display information customized for the user, and while this could occur to some extent with just interactive text, it wouldn't be nearly as effective. A good example from BBC Brasil talks about which jobs are most lacking internationally. Users can see the jobs corresponding to each country and the countries corresponding to each job. Once someone has chosen a job/country combination, there is a text box that talks about that country's immigration policies and typical wages someone could expect working there. With very simple organization, it allows users to see their options given their skillset and what countries they want to go to, and gives additional information based the job and country of interest.

`http://www.bbc.co.uk/portuguese/noticias/2013/04/130326_wanted_migrants_clickable.shtml?ocid=socialflow_facebook_brasil`

Another variant of interactive infographics allows users to interact in a way beyond just filtering information. The graphic from the Spanish newspaper *Público* examines people's opinions regarding a monarchy versus a Republic based on data collected from Twitter and this newspaper; the data show results by sex and by semi-autonomous province of Spain:



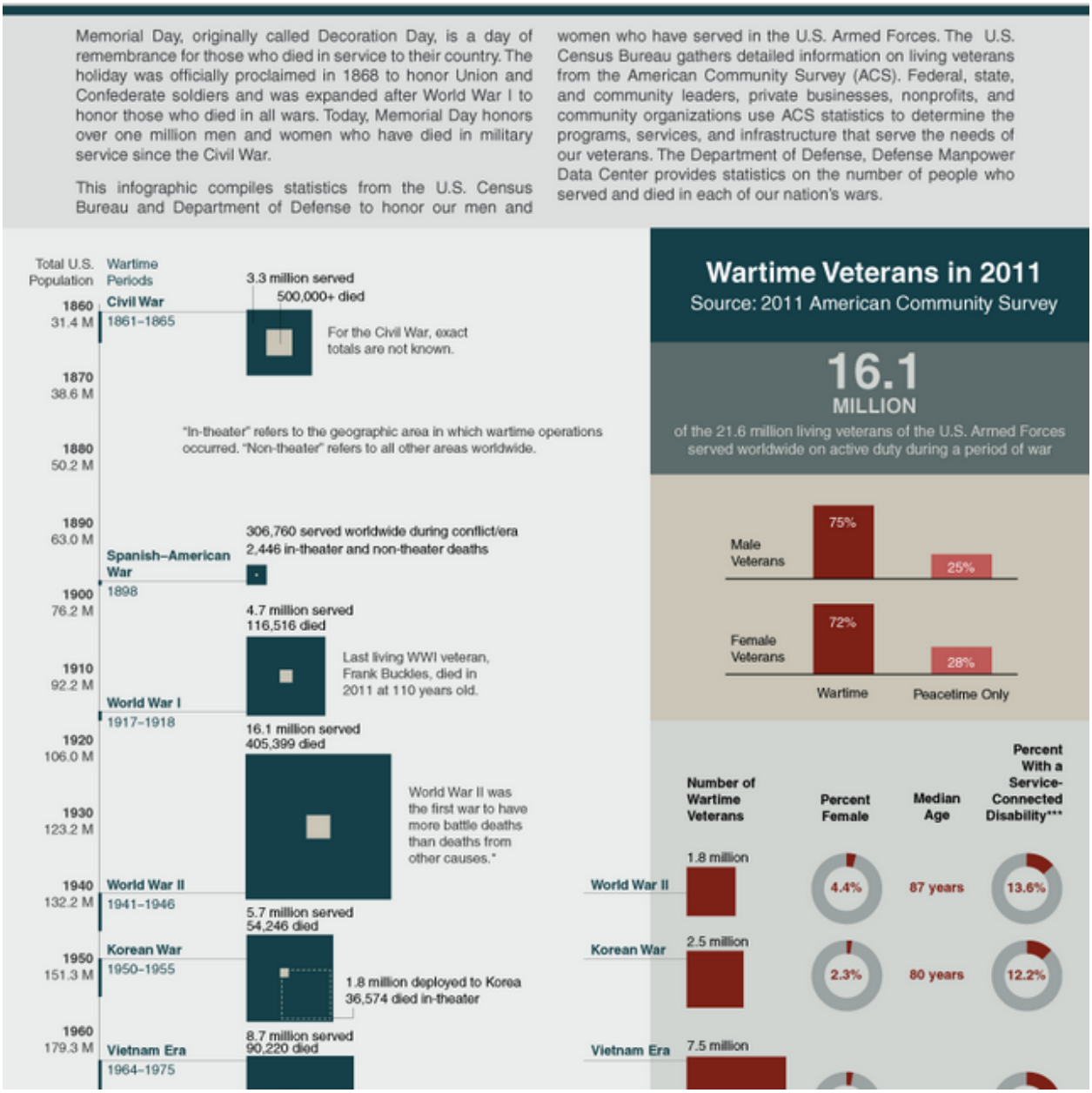
`http://www.publico.es/especial/plebiscito-virtual/?cf=publico.es`

A feature at the bottom of the page allows readers to then vote for a Republic or monarchy, along with voting on related questions like which is more democratic and to express approval or disapproval of each member of the Royal House ("Casa Real"). (When voting they also indicate their sex and region.) This information is fed into the graphic for future viewers, so they actively participate in its creation.

This infographic is definitely not without faults; in particular, none of the bars have length proportional to their actual value. (In the above picture, the "en contra" part would be nearly 40 times as long as the "en favor" part if they were proportional.) Nonetheless, overall I think it promotes a deeper understanding/analysis and generates a lot more interest than text-based alternatives.

I don't mean to imply that data visualizations are only beneficial when they're interactive, because they're not. All of the data visualizations above (except the one with pie charts) visually present information in a very effective way; this is augmented by their interactive features but ultimately based on good design principles. Nonetheless, examples of non-interactive visualizations may be useful to illustrate how they can still be useful, particularly because interactivity isn't usually an option for print versions of newspapers.

First, consider this infographic from the Washington Post this past Veteran's Day:



http://www.washingtonpost.com/blogs/answer-sheet/wp/2013/05/26/memorial-day-in-one-big-infograp

After giving a bit of context about Veteran's Day, this infographic visually shows both the number of soldiers employed and the number that died in major wars in the U.S. Because they are shown together these two numbers can be easily compared, and readers can also compare these figures for different wars preattentively (i.e. without requiring conscious effort). It's immediately obvious that more soldiers were deployed during WWII than any other war, and it's fairly easy to see quickly that more Americans died in the Civil War and WWII than were deployed in the Spanish-American War, but otherwise the comparatively small casuality rates means that deployments in

any war exceed deaths (of U.S. soldiers) from any other war listed here. Annotations give additional information (like "WWII was the first war to have more battle deaths than deaths from other causes"), and another set of graphs on the right shows demographic characteristics of surviving veterans from different wars. This infographic conveys a lot of information efficiently, makes it easy to find and compare data, and draws the reader's attention more than just text would. And importantly, this works equally well online and in print.

A second example shows how infographics can be helpful when the information presented is non-numerical This infographic from the Brazilian magazine *Superinteressante* illustrates how sex change procedures work:

Figure 4.13 "How sex change surgeries work." Renata Steffen, William Vieira, Alex Silva and Sergio Gwercman. *Superinteressante* magazine, Editora Abril (Brazil).

`http://thesocietypages.org/graphicsociology/`

It's worth noting that this infographic has a lot of information encoded as text, both labels and in explainatory paragraphs. Despite this, and despite the lack of visual encoding of numerical data, someone who doesn't speak Portuguese can still get a fair amount of information about sex change surgeries from this infographic, at least if they know what it's about (worth noting; the magazine featured a female-to-male procedure image alongside this one, which makes the context even clearer). This speaks to the visual power of the image, which is enhanced by the accompanying

text. Not only do these illustrations make the surgical process clearer, tey also make the information more engaging and more memorable. The visualizations hown above do the same thingl this one just does it especially well, and for non-numerical data.

To reiterate, I think the trend toward including more data visualizations, particularly interactive visualizations, is a positive one. Except in cases where the infographic is so badly designed that it actually obscures information, visualizations are more engaging, informative, personalizable, and data-rich than the alternative. As long as it is done by respectable newspapers and people who have a basic understanding of how to visualize data, this is a positive change (rare exceptions like the 3-D pie/line chart by NY Times notwithstanding). Data visualization should by no means totally replace standard text, but it can be an effective complement that adds a lot of information and makes it clearer, more memorable, and generally more accessible than text alone.

```
Referenced http://www.personal.ceu.hu/tex/breaking.htm
for how to do a page break in LaTeX
```

5.5 Nonparametric Statistics!

You are appointed to the role of Statistics Liaison between the social sciences faculty and the mathematics faculty at a fictional small liberal arts college. Congratulations!

Prepare a discussion piece for both faculties on the topic of parametric vs. non-parametric statistics that lays out the differences between these approaches and your thoughts on how and when these should be addressed within the statistics curriculum to best serve the needs of students in the social sciences. Use of examples encouraged.

I should preface this by saying that I have a strong bias in favor of nonparametric statistics.

The field of nonparametric statistics is a more general form of statistics that works for all types of populations. Unlike other methods, nonparametric methods do not assume anything about the underlying population in a statistical distribution, and therefore are applicable in more situations and with less information. Because they work for any data, in a sense they should just be called "statistics", while other methods should be called "normal statistics", "binomial statistics", "Poisson statistics", etc. Instead they are called nonparametric, in contrast to statistics based the normal distribution (and defined by two parameters: mean and standard deviation). Nonparametric statistics allow us to move beyond "the dogma of the normal distribution" and draw conclusions aout all types of populations (Noether 129).

There are good historical reasons for the focus on "normal statistics", mainly having to do with computing power. Nonparametric methods are highly computation-intensive, while "normal methods" are (comparatively) computation-light and, perhaps more importantly, can make use of a relatively small number of precalculated statistics tables in a way nonparametric methods can't. Additionally, normal statistical tests are robust, so that they generally work relatively well when populations only approximate a normal distribution. Think of them like heuristics: not perfect, but very efficient and good enough for most purposes. (When data actually are perfectly normally distributed they are exact, but this is actually quite uncommon. Many things are just very close approximations, and many others are approximations that *aren't* especially close.) Heuristics are important and extremely useful, but historically the statistics profession has been based almost entirely on normal statistics.

Since the development of computing and the advancement of nonparametric methods, they have become much more practical. Now that computational power is only rarely a barrier to statistical analysis, they have much more applicability than ever before (Higgins xv). This has been true since at least 1976 (when *Introduction to Statistics: Nonparametric Approach* was first published), and their usefulness has increased continuously ever since as computers have advanced. Despite the continued dominance of normal statistics, some of these have become popular and widely used, for example, Spearman's rank correlation test (an alternative to Pearson's correlation test).

Most nonparametric tests are based on permutations of data, and inferentially they test the hypothesis that a random arrangement of the data points would result in an arrangement at least as extreme as the sample more than 5% of the time. For example, when comparing two populations with normal statistics you would examine how far apart the means are and whether that difference (given the standard deviations and number of data points) provides evidence that the distributions are different. With nonparametric methods, you would look at whether one sample had a large number of data points that are higher than the data from the other sample, and whether there are sufficiently many larger points (given only the sample sizes) to conclude that the populations are different. These nonparametric methods are valid whether the distribution is normal, exponential, uniform, fractal, or anything else. They even work if the data is only ordinal, i.e. doesn't have

numbers that one can perform arithmetic on at all! (Ordinal data are extremely common in the social sciences, in Likert scales, comparisons of different treatments in experiments, things like ability that can't be easily quantified, and elsewhere.) There are parallel nonparametric tests for most normal statistical tests, including one-tailed and two-tailed tests, and many more available for different types of data with conditions less restrictive than being a normal distribution (i.e. tests can be different for ordinal vs. cardinal data, for symmetric distributions, etc.).

Because they are based on permutations rather than complex equations and proofs (and use median rather than mean as a standard measure of center), nonparametric methods are more intuitive to students and can be learned more easily. As there are variants of these tests depending on what students do know about underlying distributions, they cause students to think more deeply about the nature of their data[3] and are therefore superior pedagogically. Because they are general and don't assume conditions about distributions that are rarely strictly true, they make more sense philosophically as the default statistical test (especially with ordinal data, normal tests don't make philosophical or practical sense). Finally, because they are usually only slightly less powerful than normal tests (Noether 134), and in many cases much more accurate (especially when there are outliers), nonparametric tests make more sense practically.

Let me detail some of the applications quickly. Anything that doesn't involve a normal distribution should properly be done with nonparametric statistics; the further away from normal a distribution is, the more important this becomes. For analyzing ordinal (non-cardinal) data, whether it's Likert scales or anything else that can't be easily quantified, nonparametric methods are essential. In graphing, loess curves provide an alternative to standard linear regression for any relationship that isn't strictly linear and allow a much better analysis of the data. Nonparametric methods are generally insensitive to outliers, and medians are usually better measures of center than means when the two differ[4]. Again, nonparametric methods make much more sense as default tests than normal methods, and can sometimes even outperform them with normal data. Moreover, the types of data that do not use normal distributions are especially common in the social sciences, both because an unusually large proportion of biological and physical data are approximately normal and because much of social science data isn't easily quantified. For these reasons, I think there must be more of a curricular focus on nonparametric statistics in the math department, particularly in courses designed for social scientists.

I recommend we integrate nonparametric statistics with the statistics curriculum over a period of five years. This will give the faculty time to learn relevant nonparametric methods, allow adequate communication between math and social science faculty, and allow us to create the relevant

[3]For example, teaching nonparametric statistics directly lends itslef to asking a series of questions about the data, such as this:

1. Does this dataset involve cardinal numbers, or are values only ordinal?
2. If cardinal, are the data symmetrical?
3. If we need to take the mean, does it make more sense to use a trimmed mean to control for outliers?
4. If we're doing ANOVA, are the alternatives ordered? Do we expect both the columns and the rows to be ordered?

All of these correspond to different nonparametric tests (and different, common types of data); thinking through these questions allows students to use the most powerful test that's still valid.

[4]In a normal distribution (or any symmetrical distribution), the mean and median are the same. The most efficient way to estimate this middle value from a sample is taking the mean, and so this is typically thought of as the mean. Unfortunately this preference for means over medians as measures of center often extends to non-normal data, even though the median is more helpful in most cases (the main exception is expected value calculations).

curriculum. Unfortunately no introductory statistics textbook published after 1991 (that I know of) takes the nonparametric approach as the basis for all statistics, which I think is the place we ultimately want to be. As we transition to this curriculum we can build up resources in terms of worksheets, R code, lesson plans, and other things. Additionally, if we decide from the early years that we don't want to move to a curriculum based entirely in nonparametric statistics, we'll still have added the parts we deem most relevant/important to the curriculum.

Thank you for your time; I look forward to working with you and supporting the statistics curriculum in math and the social sciences any way I can!

5.6 Wild Card: Black Swans and Statistical Modelling

Is there a topic on which you were hoping for a question but one did not appear? Or a topic on which you wish you'd had more chance to demonstrate what you can do? Set yourself a question and answer it. You can take it from a textbook, or elsewhere, or make it up from scratch. It can be technical or discursive. It should, however, be a new question for you and not something that you've answered prior to this exam.

"Explain, in your own words, the importance of black swan events in statistical modelling, using a combination of theory and real-life examples."[5]

A black swan event is defined by two characteristics: it is hard to predict and highly influential. Difficulty in prediction can come from low probability or a flaw in the model (often it's a combination of both), but it's functionally equivalent. Because black swan events by definition cannot be predicted, risk-management mechanisms in systems are usually set up as though they don't exist, which intensifies their influence (in the context of riskmanagement, this is usually negative). A classic example is the collapse of financial markets in 2007-2008 (or any other time they collapsed; 1929 and 1987 are also quite good examples): (almost) no one predicted it, it had a huge impact across the world, and financial markets (because they all used similar risk-assessment instruments and were highly interdependant) were not set up to weather the housing bubble. In large part because this was considered so improbable, banks and financial traders were not modular and were therefore highly susceptible to systemic risk (which they could not predict); this effectively caused the whole system to experience the crisis (instead of a limited number of companies and individuals).

Black swans show the weaknesses of statistical modelling, particularly in regards to risk-management. On the one hand, black swans occur because the model is wrong; one the other hand, they occur because of random chance, in an event less likely than the arbitrarily chosen p value but necessarily possible nonetheless. More importantly, their impact is not negligible: it is in fact usually quite significant. A general property of black swans is that modern systems are ill-equiped to survive them (this effect can be reduced, but usually isn't because statistical modelling systamtically ignores the existence of black swans), so black swans are necessarily important whenever they occur. The relevant question is therefore how common they are; while specific blacks swan events are necessarily uncommon (assuming a decent model), the likelihood of *some* black swan event occurring can be quite high.

In his book *The Black Swan*, Nassim Nicholas Taleb categorizes populations as falling into two groups: mediocristan and extremistan. In mediocristan, exceptional events are still within orders of magnitude of unexceptional events. The classic example is the normal disribution: something can be a very large number of standard deviations away (e.g. 10 standard deviations, which has a probability around 10^{-24}) while still being within a couple of orders of magnitude (if we use height as an example, someone 10 standard deviations above average would still be less than double the median/mean height). In extremistan, however, events can be comparatively unexceptional (only a few standard deviations away) and still be different by orders of magnitude. What this means is that extreme events (or other samples members) will have a disproportionate, and significant, impact in extremistan, whereas in mediocristan they have virtually no influence or importance at

[5]Other things I considered asking a question about: nonparametric multiple regression, information dashboards, and designing an *interactive* infographic

all. Anything that can be described by extremistan is necessarily non-normal (though the reverse is not true; consider, for example, a Laplace distribution). As most biological and physical data are (at least approxiamtely) normal, they usually fall outside of extremistan. Most social data are are not normal, however, and many of them display extremistan properties. This can result from nonlinearities or from populations with fractal characteristics (among other things). Populations with fractal characteristics can be modelled at least somewhat, so Taleb refers to them as Gray Swans. Since that case is easier, and is also illustrative, I'll start by looking at grey swans before returning to black swans.

Unlike the normal distribution, distributions with gray swans are fractal-like in that the probability is an exponential function of the standard number of deviations. This means theat if a random set of observations at least 10 standards deviations from the mean are chosen and another random set of observations at least 1 standard deviation away are chosen, (roughly) the same proportion will be at least 11 standard deviations away in the first set as are 2 away in the second set. The probability decreases constantly, whereas in normal distributions the probability decreases at an accelerating rate. This contributes to the disproportionate influence of grey swan events as a group; they're much more common than comparably extreme events from normal distributions. (Some examples of phenomena assumed to fit these fractal distributions: frequency of use of words, magnitude of earthquakes, intensity of wars, and populations of U.S. cities (Taleb 264). Whether or not they fit perfectly fractal distributions, it's clear that they display extremistan properties; consider how common the word "the" is.) Unlike "pure" black swan events, gray swan events can be modelled somewhat, although it's difficult; the estimator (for the exponent) is biased upwards, meaning even people aware of them are likely to systematically underpredict the role of gray swans (Taleb 266). An important property of normal statistics is that estimates of their parameters are unbiased (Wooldridge's *Econometrics*, p. 52).

Black swans can't be modelled reliably, and oftentimes are more extreme. A number of system qualities and sociological effects, such as the Matthew Effect, nonlinearities, chaotic system properties, interdependencies (and dependence of behavior on predictions), and varieties of non-normality, can both destroy predictive value of models and create extremistan variations in the population. Extremistan characteristics are visible in many common social measures, like wealth (or even income), and are even more extreme in many artistic areas (e.g. a few hundred books account for the majority of total book sales). In these areas, a relatively small portion of the population is extremely influential, which makes precise prediction both more important and more difficult. To use the example of the stock market again: ten days alone account for about half of the total change of stock market prices throughout the past 50 years (Taleb 276; note that this book was published in 2006, so that claim is true for the 50 years immediately *preceeding* the most recent recession). This would never be the case if stock prices were normally distributed, just as a small number of books would never dominate the market. It is only because of non-linearities and related effects that these black swans exist and have the disproportionate importance that they do.

Black swans effectively show how much statistical significance (compared with practical significance) is overemphasized in most statistical practice. Some limitations of statistical significance have already been discussed in Daniel Kalla's seminal 2012 paper *Regressing the Corruption Perceptions Index: Don't Do It!* and won't be reiterated here (in accordance with the spirit of the last sentence in this exam question). Looking at statistical significance in terms of black swans, we see that either they occur because the model is wrong (in which case calculated statistical significance is essentially meaningless), or they occur despite low probability and have so much influence that

their low frequency isn't especially important, especially in the long run. Statistical modelling of risks commonly occurs in fragile systems; a better response is to make systems robust or antifragile to model error, meaning they weather or benefit from unexpected events (as detailed in Taleb's sequel to *The Black Swan*, *Antifragile*). Focusing on the likelihood rather than the effect of risky events makes one especially vulnerable to model errors.

Black swans are practically important to statistical modelling because they're disporportionately influential, and theoretically important because they show the weaknesses on modelling. In any distribution that displays extremistan characteristics they matter enormously. Many more phenomena do exhibit these characteristics than has been conventionally acknowledged, and as our social world become more complicated and interconnected, nonlinearities will increase and more things will develop extremistan properties (at the same time, these properties will likely intensify is cases where they already exist). The existence of black swans shows the weakness of statistical theory (at least based on the normal distribution; nonparametric statistics are pretty robust) and some of the problems with focusing on statistical rather than practical signficance. Until black swan events are incorporated more into statistical modelling (including notions like gray swans, robustness/antifragility, etc.), they will continue to undermine it's practical usefulness and its theoretical validity.

Books cited in this exam:

1. *The Black Swan* by Nassim Nicholas Taleb
2. *Econometrics: India Edition* by Jeffrey M. Wooldridge
3. *Introduction to Modern Nonparametric Statistics* by James J. Higgins
4. *Introduction to Statistics: The Nonparametric Way* by Gottfried E. Noether

Chapter 6

Synthetic Thinking: A Manifesto (a.k.a. Everything is Awesome!)

Abstract

This paper is a manifesto for synthetic thinking, defined as the type of thinking that focuses on finding merit in and supporting ideas. I start with criticism of critical thinking, both in general and specifically in the context of statistics and feminist critical theory. I then make a case for the merits of synthetic thinking, drawing on literature dealing with conflict and other examples as well as specifically addressing synthetic thinking in statistics and feminist critical theory. Finally, I describe a model for determining truth from contradictory claims using synthetic thinking and conclude by discussing its relevance to natural science.

Introduction

Critical thinking, commonly considered an essential skill from academics, has at least two man distinct sense. On the one hand, "critical" can mean "crucial", in which case critical thinking typically refers to the ability to analyze a situation. On the other hand, "critical" means "criticizing", and this skill refers to the ability to find flaws in arguments. These two senses are often conflated, leading to the idea that being able to find flaws in argument or conclusions is *the* essential academic skill. Here I take the opposite approach. I believe that the ability to find merits in different arguments, here referred to as synthetic thinking, is more important and more useful in all aspects of life, and that synthetic thinking should be prioritized over critical thinking. This paper sets out to show the superiority of synthetic thinking, drawing from a wide variety of examples inside and outside academics.

This paper is organized into three main sections. The first section, *Everything Wrong with Critical Thinking*, sets out to show the flaws inherent or common in critical thinking after briefly describing the history of critical thinking and its ties to the scientific method. I illustrate some of these by looking at examples in statistics and feminist critical theory, both because these two areas are illustrative of broader trends and because they're directly relevant to most aspects of life. The second section, *Why Synthetic Thinking is the Best Thing Ever*, describes the merits of synthetic thinking and how essential it is to success in many areas of life. I draw primarily from examples about handling conflict at first here, before addressing statistics and feminist critical

theory again from a synthetic approach. In the final section, *Deciding Between Two Claims with Synthetic Thinking*, I describe a model for determining truth from two contradictory claims, using synthetic thinking (more than critical thinking) to choose one claim over the other in the first case and using synthetic thinking to integrate the two claims in the second case. To conclude, I refer back to the scientific model described as the basis for critical thinking and use it to illustrate the roles of synthetic and critical thinking when the two are used together.

The main idea of this paper is drawn from the Appendix of Peter Elbow's *Writing Without Teahcers* [54, *Appendix*], in which he defines the doubting game and the believing game. To anyone familiar with his work these terms can be used interchangeably with "critical thinking" and "synthetic thinking" throughout this paper. I use criticaland snthetic thinking because critical thinking is a well-established term in education literature, and because of additional meanings of synthetic thinking I'll explain in that section later.

Throughout the paper, there may be things the reader disagrees with or times when I oversimplify an important issue. Given the topic of the paper I encourage the reader to try to see past the flaws, and least temporarily, and make sure to understand the merits of the argument before rejecting it entirely. On that note, enjoy!

6.1 Everything Wrong with Critical Thinking

Critical thinking is used throughout this paper in the sense of criticizing and doubting conclusions, institutions, and society. Historically, critical thinking seems to be a result of the Socratic Method of questioning all conclusions, later described as "doubt everything" by Descartes. It has since turned into the scientific method, which aims to falsify all claims in science and accepts those which it is unable to falsify after many experiments. In the scientific method, hypotheses are *required* to be falsifiable for them to scientifically testable, and thus a major challenge for scientists is designing good experiments that could falsify one hypothesis. With Socrates's influence in Western philosophy and the incredible advances of natural sciences in the past centuries (and consequent apparent centrality to the natural sciences in modern life), critical thinking has achieved dominance in academia, and as a result has become a major influence in how people think in all walks of life.

Essentially, the idea behind critical thinking is that when flaws are identified, they can consequently be fixed, leading to a better institution or knowledge system. Critical thinking encourages perfectionism, where perfection is understood as flawlessness. By always working to identify (and thereby eliminate) flaws, critical thinking aims to remove weaknesses or imperfections. This means that, in order to be useful, it must constantly identify these flaws. A system where flaws cannot be identified cannot be improved by critical thinking, and one that exhibits flaws cannot be fully accepted. To demonstrate some of the flaws of using a strictly ctitical thinking approach, I show what it means for statistics and feminist critical theory. I'll start with statistics, primarily because it's more directly related to the sciences and therefore more susceptible to the influence of the scientific method.

6.1.1 Statistics and Critical Thinking

There are essentially two types of statistics where hypotheses are tested: inerential statistics, which tests whether a given statistic has a certain value or whether two populations are equal; and correlational statistics, which tests whether two different variables are related to each other. In both of these methods, the data are subject to statistical variation, and therefore any conclusions

are uncertain. Statistics deals with this through the concept of statistical significance, with is essentially a measure of how (un)likely a false correlation is to appear due to random statistical variation. The statistical significance value that is necessary to accept a hypothesis is essentially arbitrary; by tradition, statisticians usually use value of 5%, or sometimes 1%, as the cutoff point (meaning that such a result would occur as a result of random statistical variation less than 5% or less than 1% of the time). In the best of cases, this means that up to 5% of "statistically significant" results will be wrong. In most cases the number is almost certainly higher, because statistical modelling makes heavy use of assumptions. Any sort of error in the model can lead to a much higher uncertainty, which means that the critical thinking stategy of accepting conclusions that haven't been shown to be wrong leads to a large number of false positives.

Statistical hypotheses are based on assumptions about the data, of which three are particularly important. First, most modelling assumes that the population is sampled randomly, with each one chosen independent of the other ones. Second, it assumes that the variables under consideration are actually the variables the modeller wants to measure. Finally, most analysis assumes that the data follow a normal distribution (or bell curve). In practice, usually none of these asumptions is strictly met, and it's relatively rare for all three of them to be met completely (which still doesn't fix all the problems). The problems involved in meeting these three conditions are discussed below.

Almost any way of sampling a population falls short of being random, which can affect the results. For example, calling random phone numbers restricts the population to people who have phones, and oversamples people who have multiple phones or who use phones frequently. (Also, as with any verbal survey method, it requires that the people being sampled participate and excludes people who don't respond to surveys.) This has the potential to distort results so much that an example of a major phone survey totally mispredicting the presidential election is used in many introductory statistics textbooks to make this point. Even with a truly random survey, anything requiring followup will probably result in some people being excluded from the data set, which is likely to bias conclusions. Even if the military draft was random[1], a correlation between fighting in the army and lower life expectancy may not be valid because some people may have chosen to enroll in college, run away to Canada, or take other steps to avoid fighting in the army, meaning the group actually in the army is not random. As similar problems apply to most other possible statistical models, it is extremely difficult to get a sample that can truly be called random.

Even when that occurs, statisticians run into major problems due to variable specification. Consider a variable like intelligence. It's unclear exactly how such a thing could be defined and measured, which means that anyone trying to statistically test conclusions about intelligence must use a proxy variable. Usually they use IQ, GPA, or some standardized test score as their measure, and then relate IQ score to income (and claim that more intelligent people earn more). Because test scores are not the same thing as intelligence, something can cause a difference in test scores without effecting intelligence. Variables such as fluency in English, motivation, confidence, stereotype threat, hunger, hours of study, and other factors are all likely to affect test scores even as they leave intelligence unchanged (which is part of why retaking a test usually results in a different, albeit similar, score). In this case the choice of income could also be criticized (Why not wealth? Why not net income after paying interest on student loans? If the study is cross-cultural, is income measured in dollar amount or purchasing power (and why not the other choice instead)?), creating more opportunities for error and making the analysis much less reliable than "95% certain".

[1]There's evidence that it was heavily biased toward later months, meaning birth month may have been a confounding variable, so strictly speaking this conclusion would be statistically invalid anyway [136].

There are many, many other common statistical errors, as deatiled in this footnote.[2] Many of these, like the ones above, are generally acknowledged by statisticians and used to criticize other people's use of statistics. The assumption that data are normally distributed is noteworthy because it's usually *not* acknowledged by statisticians, but still leads to (often drastic) model error. Nassim Nicholas Taleb documents many of these in his book *The Black Swan*, including creating a fractal distribution and noting some of the things it describes better than the normal distribution [157, p. 262-273]. In general, a physical variable like height will be roughly normal (although not perfectly normal, which still somewhat invalidates the model), while most social variables will not be. (Sample averages are also close to normal, for any distribution with enough data.) While most statistical tests still approximately work when the population distribution isn't perfectly normal, this does increase the uncertainty above the claimed value and sometimes give wrong conclusions.

In theory, a greater understanding of statistics should increase someone's understanding of the world. In practice, understanding statistics often means that people can reject statistical evidence whose claims they disagree with for one of the reasons given above or a different statistical criticism. Even worse, ambiguity and statistical variation in evidence mean that they can usually find statistical evidence supporting even false beliefs, and perhaps convince other people of false beliefs also. "Lying" with statistics is therefore very easy (second only to lying *without* statistics), intentionally or otherwise. Because the field of statistics is so strongly focused on being able to reject conclusions, people can adopt a belief, criticize or deny anything that contradicts it, and become increasingly sure of that belief as contradicting evidence is presented. While I certainly think statistics can be useful (especially when used synthetically!), they can also quite easily lead to intellectual intransigence and make it harder to find the truth. A good example of this comes out of the discourse surrounding climate change, from a video called *C(Lie)mate #3 - What Doesn't Make the News* (and some of the author's earlier videos about the same thing).[3]

Denying Climate Change – with Statistics!

In this video, Suspicious0bservers (S0) attempts to statistically refute the mainstream scientific idea of human caused climate change. He points to the correlation between CO2 levels and temperatures that lasted for over 300,000 years, and then stopped being true after the start of the Industrial Revolution [1, 0:18]. Since humans only started contributing significant amounts of CO2 to the atmosphere in the industrial revolution (and based off scholarship mentioned in one of his previous videos that describes how changes in temperature might cause changes in CO2 level instead of the other direction [2, 2:15]), he concludes that the human contribution to climate change through CO2 emissions doesn't seem to be very significant. S0 points to a strong correlation between solar activity and temperatures [1, 0:39], noting that 4 of the 5 highest peaks of solar activity since the ice

[2]Some other common methods include bias in research procedure (e.g. experimenter/subject expectancy); bias in variable measurement (e.g. standard white English is considered more correct than equally complex black English in a cognitive test); exclusion of part of the population (see analysis of Pettit elsewhere in my Plan); random chance; low pre-experiment probability (esp. in case of philosophical disagreement) [91]; publication bias (non-results are usually not published); arguments of which statistical test is appropriate; recording errors; lying (on surveys); biased questions (on surveys); unique circumstances that prevent something from generalizing (or it being a very limited scale); and other reasons too numerous to mention. Also, for inferential statistics one can almost always make the 'correlation doesn't imply causation' argument, and it's usually possible to make the reverse 'causation does not imply correlation' argument (e.g. people who go to doctors frequently might not be healthier, despite the fact that going to the doctor is probably good for the health).

[3]I also draw from Glenn Beck's *An Inconvenient Book*, but for the sake of brevity I keep that discussion to the footnotes.

age in the 1600s have occurred in the past 70 years [2, 5:04][4]. He notes that the earth is long overdue for a shift in magnetic poles [1, 2:43] and that the beginning of a rapid shift has been observed for the past decade. This corresponds to a weakening of the magnetic shield to its lowest level in at least 400 years [1, 2:16], making the earth much more vulnerable to changes in solar activity. Together, S0 argues, the changes in magnetic field and in solar activity best explain climate change. This conclusion, well-supported with statistics, goes against the mainstream climatological scientific community and demonstrates intellectual intransigence being supported by statistics (either by S0, if you believe in climate change, or by mainstream scientists if you don't).

S0 is able to further strengthen his claim by looking at climate projections from the IPCC (Intergovernmental Panel on Climate Change) and other organizations and noting how rising temperatures have fallen short in almost all these cases. He notes that "predictions were on average three to four times what occurred in the real world" [1, 3:54][5] and shows a number of news articles (from very diverse sources) noting how the IPCC has had trouble explaining the lack of temperature increase in the past 15 years and why its predictions were so low.[6] [1, 3:15-4:23] S0 concludes by noting that solar activity is dropping rapidly to the point that some scientists are projecting the (at least temporary) end of sunspots by 2015 [1, 5:45 and 6:15], meaning we might be about to start another ice age [1, 6:38 and 6:59]. In other words, statistical evidence shows that we shouldn't be worried about global warming, and perhaps we should worry about the reverse.[7] Overall, S0 manages to provide a significant amount of support for his doubt of climate change, all using statistics.

It should be clear that critical thinking in statistics does not allow one to decide whether or not climate change is caused by humans. S0 is able to use detailed, thorough statistical analyses to support his contention, in spite of all of the evidence by mainstream climatologists to the contrary (equivalently, mainstream climatologists believe that humans are responsible for a significant part of climate change in spite of all the evidence from S0, Beck, and others to the contrary). Since climate change is currently one of the most highly researched areas of science (as a result of its important social implications), it's clear that this is not due to a lack of research: in fact, the depth of research might be working to *strengthen* the conviction of each side, because there's so

[4]In *An Inconvenient Book*, Beck shows the strong fit between solar activity and temperature from 1880-2000 while noting that in this time period, CO2 only rose rapidly after temperature (and solar activity) increased [16, p. 6]. Beck's graphs are more limited in time scale, which allows them to show recent trends more effectively (although it also calls their long-term validity into question).

[5]Glenn Beck does this a bit also, although he mainly picks on Al Gore (as a political pundit, Beck is pretty partisan, hence the title of his book). In addition to criticizing Gore's use of scare tactics in *An Inconvenient Truth*, Beck notes that "Gore's apocalyptic...vision of the future was based on a water level rise that is about 17 times higher than what the [UN IPCC] is expecting. In fact, it's even five times higher than the *high* end of the *worst-case scenario* painted by scientists who thought the UN report was too conservative" (emphasis original) [16, p. 5]

[6]One of the news articles mentioned in the report notes that less than 2% of simulations predicted a rise in temperature as low as it has been for the past 15 years, or "[i]n other words, over 98 percent of forecasts show CO2 emissions as high as we have had in recent years leading to more of a temperature increase" [159]. This means that either the simulations are fundamentally flawed, or it just so happens that the standard 95% certainty measure in statistics didn't work here. Either way, this effectively illustrates how statisticians can't be certain of any of their conclusions.

[7]Beck ends his chapter similarly, noting a sustained decline in number of deaths by extreme weather from 1915 to 2004 despite a growing population and a massive increase in CO2 (the absolute decline is about 96%, while the per capita decline is about 98.7%), which seems like reason to be optimistic [16, p. 20-1]. He earlier made mention of the 'global cooling' predictions coming from climatology in the 1970s and similar trends before: "The *New York Times* wrote scary articles about cooling in 1924, warming in 1935, back to cooling in 1975, and, of course, warming again today...*Time* did essentially the same thing..." [16, p. 15].

much diverse evidence supporting them. The evidence is sufficiently ambiguous that the critical thinking inherent in statistics allows each side to reject the arguments and conclusions of the other. An important effect of this is that many people also reject good institutions as intellectual authorities (e.g., S0 and others no longer think IPCC is reliable and will likely distrust its models and conclusions in the future). While statistical criticism *can* lead to rejecting good institutions or programs, feminist critical theory does so on a regular basis and will provide a more illustrative example.

6.1.2 Feminism and Critical Thinking

Critical theory is "a social theory oriented toward critiquing and changing society as a whole, in contrast to traditional theory oriented only to understanding or explaining it. Critical theories aim to dig beneath the surface of social life and uncover the assumptions that keep us from a full and true understanding of how the world works." While a case could be made for this definition using critical in the sense of 'crucial', in practice the discipline is very much focused on criticizing. Most feminist theory is heavily based in critical theory, with the result being that feminism is highly critical of most social institutions. (For convenience I shall refer to feminist critical theory as feminism from this point forward.) Largely as a result of this, many people are turned off by feminism. Some of the standard cultural images of feminism, bra-burning and trying to live completely without men, reflect people's perception of feminism's critical nature; in this way the perceived basis of feminism in critical thinking is harmful to the movement. I contend that the dominant practice of critical thinking is as well.

Since the women's movement in the 60's and 70's and consequent development of feminist critical theory, nearly every major institution in the U.S. has been criticized by feminism for various reasons[8], including feminism itself. Until recently feminism has been an unselfconsciously white, middle-class movement. Now it's *extremely* conscious of that fact, probably more so than any other major movement in history. Despite attempts to be more inclusive, feminism remains primarily a middle-class, white women's movement, both because of oversight that results in other groups being excluded and because many other groups find feminism unappealing. To the extent that it's unappealing, this is due priamrily to its history [83, p. 41-3] and its critical nature. As a result, most people in the U.S. are unsympathetic to feminism (roughly 20% of the population identifies as feminist) even though the vast majority are (at least nominally) in favor of equal rights [156] and feminism literally means "the doctrine advocating social, political, and all other rights of women equal to those of men." [49] Aside from changing its history[9], the feminist movement can only make

[8]Some of the most common feminist criticisms are that an institution is prejudiced on the basis of sex, race, class, religion, gender identity, sexual orientation, ablebodiedness, looks, size, ethnicity, and many other dimensions. The prejudice can be in the form of prejudiced ideologies or discriminatory effects for different groups (even if unintentional), can refer to the present tense or the past tense, and can be in terms of who a group or person associates with. Talking about or acknowledging race, gender, class, disability, etc. can be criticized (because it's not 'colorblind'), and ignoring it can as well. In addition, anything that can be read as silencing any group is open to harsh criticism (which commonly includes calls for civility, or attempts to have a less heated, antagonistic argument/discussion), and similarly for things that can be viewed as victim-blaming (which includes anything that treats people's feelings as invalid). Not being angry about an issue, or disagreeing about its importance, can also be cause for blame. Finally, anything that does not go out of its way to include everyone (and succeed) is likely to be criticized. (There are many, many more, but these are some of the most common criticisms.)

[9]Actually, the feminist movement is one of the most historically revisionist movements in effect today, as it constantly tries to highlight examples of important women (particularly women of color) in history. It even invented the term "herstory" to try to correct for the perceived male bias in "history". None of this is said as a criticism,

progress here by reducing it's reliance on critical thinking.

A good example of this is the common phrase "check your privilege". Feminists often use this phrase whenever someone makes a claim they disagree with, with the implication that there's a good chance they wouldn't make the claim if they weren't a member of a privileged group. Essentially, though, this functions as a silencing tactic for opinions they disagree with. Anyone who is privileged (which includes probably everyone in some way or another) can be silenced this way, which among other things can make them feel unwelcome in the movement. Even feminists are severely handicapped by this; groups of white women often express the feeling that they can't talk about race effectively because there's no one there to share their perspective as a person of color, which severely limits what they can do. To avoid speaking for people of color, they may avoid talking about race at all, which is counter-productive. Even if they do, people are often afraid to express opinions, limiting the amount of real dialogue that can happen. While consciousness-raising can certainly be useful, it is less so when it shuts down discussion and is entirely negativistic. Negative, intense criticisms of individuals can be very off-putting.

Criticism of institutions and ideas can be even more intense. Because nearly anything can be criticized on some feminist ground, any institution that someone dislikes can be rejected on ostensibly feminist grounds. It becomes exceedingly easy to justify one's private prejudices for this reason, even while taking the moral high ground. This leads to the same problems as statistics, although in feminism it's arguably even more extreme. Regardless, in addition to allowing people to reinforce their own prejudices, the push to dissociate from and criticize institutions means that feminists often fail to see the positive aspects they have. Critical thinking does not help them find positive institutions that, overall, they like, nor does it help them build other institutions with qualities that they'll like. (In terms of failing to find institutions they like this isn't necessarily a problem for many radical feminists, because they want to dismantle capitalism and all the associated institutions. Even there, though, they'll have trouble constructing positive new institutions.) Feminists have become extremely good at criticizing different systems but really struggle in trying to build new systems.

Indeed, by focusing only on these negative aspects actively prevents feminism from being unified, because different feminists contradict each other. While there are any number of examples available, perhaps one of the most striking concerns the concept of "trigger by void" from a 2009 blog post [110]. Essentially, in this post the author talked about how the absence of sexual assault in the entirety of the show of Lost draws attention to sexual assault for her, because the threat of it would definitely be present and so acting as though it weren't is unnatural. Featuring sexual assault in the show would be triggering to some people and (to use the feminist phrase) highly problematic. The author acknowledges this, and doesn't know what alternatives are possibly available. In watching the show, she is actively looking for something bad (even if unconsciously) so that when it doesn't appear, the show feels off and merits criticism. By focusing on critical over synthetic thinking, she makes it so there's no way out of this situation (at least not a clear one). This is an extreme example[10], but it follows directly from the exclusive focus on critical thinking.

---

just to highlight the fact that historiographic changing can only be so effective, and therefore the movement should consider focusing more on reducing its reliance on critical thinking.

[10]Discourse is a concept from Foucault that is often used in feminist epistemology to make general claims using individual examples. Essentially, the idea that a given claim only makes sense for people to even think in certain environments. Here my claim is that a concept like "Trigger by Void" would only appear on a reasonably popular feminist blog if the movement as a whole was highly focused on problematizing media, which means that this example is relevant despite being extreme and anecdotal. (I consider the blog reasonably popular because it has enough of a

A more typical example is something along the line of *I Don't Watch Orange is the New Black*, an article in which Alison Samuels talks about why she refuses to watch the show [141]. Her argument is that Hollywood has avoided portraying racial diversity in other shows where it definitely should feature (for example, *Friends*, set in NYC). Because it's only willing to do so in a show about a prison, and therefore in a stereotypical, negative role, Samuels does not want to support the show. While there is merit to the criticism, it's unclear what a show about prisons should do in this circumstance since using an all-white cast has its own problems. Both *Orange* and *The Wire* (another show Samuels criticizes in the article, for the same reason) address some deep race and class issues and certainly have at least the potential to be positive. This criticism leaves producers with no place to go, and is therefore extremely disempowering. There are numerous examples like this, and relatively few models of what producers can do in situations like this (except "avoid"). This requires synthetic thinking; doing so otherwise is essentially impossible.

To summarize, critical thinking is a dominant emphasis in academia and a major part of how people think in society more broadly. Statistics and feminism directly relate to most parts of acadmeia and society, while also illustrating the broader trend. A dominant focus on criticla thinking not only often fails to find the truth, but often actively works against this by allowing people to not consider evidence or viewpoints that disagree with theirs. It also causes people to reject good institutions without giving them the tools to build better, more desirable instituions. Because my training is mainly in critical thinking I have spent most of the paper so far criticizng; from this point forward I will be promoting synthetic thinking, and a framework that combines synthetic and critical thinking.

6.2 Why Synthetic Thinking is the Best Thing Ever

Synthetic thinking can be defined as "the combination of ideas into a complex whole", or in some cases as "[t]he combination of thesis and antithesis in the Hegelian dialectical process whereby a new and higher level of truth is produced." [85]. This means that it works by the opposite process of critical thinking; whereas in critical thinking the goal is to gather evidence to attack a conclusion, in synthetic thinking evidence is brought together to support a conclusion. Equivalently, the process of critical thinking involved finding flaws in an argument or reasons to doubt it, while synthetic thinking involves finding merits in an argument or reasons to believe it (and similarly for supporting or oppossing institutions). Given the centrality of critical thinking to academia, it could also be seen as representing academic intelligence, whereas synthetic thinking represents (among other things) creativity. Especially in a rapidly-changing world such as our own, synthetic thinking represents a vital skill and deserves much more emphasis than it's getting in education.

Creativity is a skill that is becoming increasingly important, but schools and teachers really struggle to teach it. I think this is for two main reasons. First, there is a lack of institutional support for creativity; the amount of time and infrastructure devoted to taking and preparing for standardized tests drastically curtails the ability of schools to teach creativity. Second, there is a lack of knowledge on the teachers' end. In part because schools have historically been ineffective at teaching creativity, teachers don't have strong models of how to teach it effectively. What this means is that they don't have the structures in place to teach creativity. Instead, the common approach is encouraging people to think "outside the box", meaning to abandon their preconceived ideas

presence that there's a blog devoted to criticizing it, http://stfushakesville.tumblr.com/)

and biases and thereby produce a creative result. This is a very critical thinking-style approach, and while it does sometimes produce creative individuals and ideas it fails to promote creativity systematically (in part because, despite attempts to make the environment nonjudgemental, people fear criticism) [25, p. 30-2]. Because it focuses on problematizing current ways of thinking and doesn't provide a structure for thinking creatively, this method is often ineffective.

Fortunately, there is at least one structure in place specifically to increase creativity. Systematic Inventive Thinking, or SIT, takes the approach of using a structure to systematically be creative. Regardless of the merit of this program specifically, it's philosophically very appealing because it promises a route that can reliably produce creative results [25, p. 8-9]. SIT, fortunately, is definitely teachable , and it or a similar model could fairly easily be incorporated into schools. It's been successful so far in a number of businesses [25, p. 3-4], although regardless of whether SIT specifically proves effective for schools in general, it's clear that a structure for teaching creativity is needed. I claim that synthetic thinking is much better for building systems and alternatives than critical thinking (almost by definition) and that it should be taught alongside critical thinking in schools. Before that, though, I want to make the case that synthetic thinking alone can be extremely powerful in a variety of fields, including conflict resolution, nonviolent resistance of dictatorship, statistics, and feminism.

6.2.1 Conflict Transformation

Conflict resolution is an essential skill for all spects of life, because conflict pervades all parts of life. People struggle with conflict resolution for many reasons, the most prominent of which is the thirteen barriers to communication. The first twelve barriers barriers, classified under broader categories of judging, sending solutions, and avoiding concerns, are all wrapped up in critical thinking; the thirteenth, pointing out that the other person is using a barrier to communication, is even more so [22, p. 17; 25]. As most people who have been in heated arguments can probably attest, criticizing the other person is unlikely to help, and extremely likely to escalte, a conflict. Criticisms of policies or institutions, while separate from the barriers ot communication, are also likely to start or escalate a conflict, for the same reasons. Critical thinking approaches to conflict resolution are spectacularly unsuccessful, creating and exacerbating problems much more than they ameliorate or solve them. Hence, there is a need for synthetic thinking.

While there are many important skills involved in conflict resolution, probably the most important two are the ability to understand another viewpoint and the ability to develop a solution that meets the needs of both parties. Understanding other viewpoints, particularly viewoints that contradict one's own, is what synthetic thinking ("the believing game") is all about. Moreover, in conflict resolution there are many specific concrete skills that help one do so (in particular, reflective listening). Actively building on these skills, and subsequently using them to understand the other person's viewpoint, is the most successful (and synthetic) approach. The other main skill in conflict resolution, finding solution that meet people's needs, directly involves comining two (ostensibly) contradictory viewpoints. While critical thinking by definition involves finding flaws in (the other person's) solutions, synthetic thinking by definition involves finding the merits of solutions from both sides and combining their positive attributes in some way. Any time someone wants to resolve a conflict, they must turn to synthetic thinking for it to be successful.

The biggest objection to my formulation here is that sometimes people *do not* want to mutually resolve conflicts, except by beating the other side. While this rarely desirable in the long-term in most walks of life, there are some circumstances where this may be legitimate, most notably civil

rights issues. In these cases people may think that legal equality is the only solution that meets their needs (which is probably true for civil rights issues)[11], meaning other people's solutions are not acceptable alternatives. The people want a revolution!

First, it is worth examining the general process of effecting change in a society. In *Rules for Radicals*, Saul Alinsky notes that the first step to effecting change is always to use the pre-existing, legitimate political channels available [4, p. *xix*]. Even if one doesn't believe they will work, trying them first will give much more ideological support for any other type of action later, legal or otherwise.[12] If using exisiting systems proves ineffective, Alinsky advises using these systems in different way. One of my favorite examples from his book is protesting at banks by having large numbers of poor people open accounts, and then close them the next day. With enough people doing this, the amount of resources the banks need to devote to processing applications means they're effectively handicapped [4, p. 162]. Because people are taking actions consistent with the banks normal procedures rather than trying to undermine banking a different way, the bank cannot counter this form of protest except by not allowing people to open accounts, which effectively accomplishes the same goals. While banks can counter critical protests by investing more actively in their normal processes, they can counter these synthetic protests only by designing a new system. As Salinsky notes, any goal that falls short of destroying the system of government entirely will have the highest chance of success if protesters make use of existing institutions and legitimate channels of protestnas much as possible. This means that the only time conflict could possibly be resolved better with critical thinking is in the case of a complete, radical revolution.

The best example of a nation requiring a total revolution is a dictatorship; the change to democracy will require a complete restructuring of governance structures, rulers, and institutions to enforce the law. This is also a good example because, in the seminal book *From Dictatorship to Democracy*, Gene Sharp explicitly warns against negotiation, on the basis of the two sides having irreconciliable goals (and because of the danger of minor concessions undermining the force of a resistence movement). Despite this, Sharp notes that nonviolent resistence is the most effective way to overthrow a dictatorship and establish a democracy, for a few main reasons. First, he notes the importance of having a vision and a strategic plan, with all actions forming part of that plan and working toward that vision. This means that even negativistic acts, like strikes and symbolic criticisms, are in fact directed to a positive end, and designed to cause ideological conversion rather than directly destroying a system. (In *Prison Notes*, Barbara Deming notes that nonviolent resistence would not be effective if it relied exclusively on noncooperation, but can be extremely effective because it also makes use of ideological conversion.[13]) Second, Sharp advocates building a parallel government along side the dictatorship, both because this greatly increases the likelihood

[11]Gene Sharp, described more below, said in *The Politics of Nonviolent Action* that "some conflicts do not yield to compromise and can be resolved only through struggle", continuing that "[r]egular institutional procedures are rarely available [to solve] conflicts that, in one way or another, involve the fundamental principles of a society, of independence, of self-respect, or of a people's capacity to determine their own future." [145, p. *xv*]

[12]Also, of course, legal channels very frequently *do* work to resolve civil rights issues. Most importantly, unless a government is going to be toppled entirely, any change must come through legal channels. A good example of this currently is the fight over gay marriage, where gays are earning marriage equality in different states through a combination of legislative decision, (state) Supreme Court decisions, and voter referenda.

[13]If we relied, in our struggle, simply on noncooperation...it *would* be absurd of us to hope to move them...The point is, we don't rely simply on noncooperation, and we don't rely on it as a bludgeon. I have called nonviolent action a dramatic technique...[i]t is, to be sure, a kind of force; but we are not trying with it to force those opposing us to their knees; we are trying to force them to look at a situation in a new way." [48, p. 71-2] Everything these activists did in nonviolent resistence was dedicated toward creating another point of view for the other person, with the hope of ideological conversion (partial or total).

of transitioning to a peaceful democracy when the current government falls and because it actively shifts powers away from the current regime and gives the democratic movement more resources to resist effectively. A third reason nonviolent resistence is more effective[14], noted by Erica Chenoweth and Maria Stephan in *Why Civil Resistence Works*, is that nonviolent resistence more easily allows partial victories to occur (example in footnote[15]) [34, p. 9], while violent resistence can usually only be successful by defeating the current regime entirely. Fourth, nonviolent resistence can be more inclusive and allows many more people to participate in ways meaningful than violent resistence]footnoteThis is especially true of people *within* the ruling administration, who are more likely to sympathize with nonviolent revolutions and able to participate with less risk to themselves [34, p. 47]. It is worth comparing nonviolent resistence to *The Art of War*: while much of the advice is unhelpful (in particular, secrecy and suprise attacks are often counter-productive in nonviolent actions [145, p. 62]), even in violent struggle "[s]ubjugating the enemy...without fighting is the pinnacle of excellence." [154, p. *xix*] Loyalty shifts, while rarely sufficient in themselves, are utterly essential to nonviolent warfare., and appeal to more people because they may be less dangerous. These theoretical reasons were born out in practice also, as Chenoweth and Stephan show how over the past century nonviolent resistence has consistently and dramatically outperformed violent resistence statistically in the most difficult of circumstances [34, p. 8, 9, 74].

Conflict transformation requires synthetic thinking at all levels, from negotiated resolution to forceful revolution. At the most basic level level, critical thinking inhibits effective communication and both causes and exacerbates conflicts. Synthetic thinking, both in understanding the other person's ideas and in finding a mutually beneficial solution, is able to resolve conflicts between two parties more effectively than any other strategy. Social change is most effective when it happens through legitimate channels, first used legitimately and then (if necessary) used for new purposes to promote the revolution. In the extreme case of trying to end a dictatorship and establish a democracy, nonviolent resistence based in synthetic thinking still outperforms violent resistence in all fronts: number of participants, range of strategies used, likelihood of success, ability to establish a peaceful regime after the dictatorship falls, and possibility of partial success. No matter what type of conflict transformation is intended, synthetic thinking is the most effective way to accomplish one's goals.

[14]Sharp also explicitly warns against violent resistance, in part because the dictatorship always has the advantage in ability to exercise violence, whereas nonviolent people on the side of democracy have ideological advantages [144, p. 32, 31]. Essentially, rebels should play to their group strengths instead of just trying to attack the dictatorship violently.

[15]Catalonia under Franco is a fascinating example of this, because it benefitted from all three main methods of nonviolent resistence (non-participation, symbolic resistence, and nonviolent intervention (specifically, parallel government)). As the economy started to grow in the 1950s and 60s, there were major strikes in Barcelona that led to loosening of government repression. Catalonia began to rebuild the arts in Catalan (the native language that Franco had repressed) in a cultural movement called *La Nova Cançó* ("The New Song"), both to resist Franco and to preserve the language. In 1971 they established The Assembly of Catalonia and a declaration of rights, including political and economic power. When Franco died in 1975, Catalonia was able to transition to being a semi-autonomous region, with significant political and economic independence and well-established linguistic rights accepted by Spain [67]. The partial victories of the movement were important before the dictatorship fell because they gave people more freedom and access to Catalonian culture, and they were important after the dictatorship fell for unrelated reasons because it helped Catalonia transition to a strong, free region.

6.2.2 Synthetic Thinking in Other Arenas

As I contend that synthetic thinking is superior to critical thinking in almost all situations, the following is a list of examples showing its superiority. After these brief examples, I will return to statistics and feminism to show how synthetic thinking is better in those specific contexts. The reader may skip this list with no impact on the rest of the essay; it's merely intended to show how synthetic thinking is applied to different areas and make my central claim more convincing through examples.

With that said, synthetic thinking is essential to understanding such diverse areas as the following:

- **Anthropology**: The key to success in anthropology is understanding why a given culture has the traditions it has and seeing what works about them before criticizing. Anthropology has progressed as a field only when scholars have understood seemingly odd traditions, instead of dismissing cultures as savage and uncivilized. Most often this happens through participant-observation, where an anthropologists both actively participate in the culture and observe it to see how it works.

- **Organ Donation**: I'll talk about assumed consent when I discuss sociocracy and feminism, but behavioral economists have consistently noted that an opt-out system (where people donate their organs when they die) results in many more organ donations than an opt-in system (where people must actively choose to become donors) while not infringing on civil rights because it still allows people a full choice [163]. The effectiveness of positive defaults is a major theme in behavioral economics generally.[162, p. 83]

- **Writing**: The idea of the believing game comes from Elbow's *Writing Without Teachers* and was the inspiration for this paper. Effectively, Elbow argues that most people who have trouble with writing do so because they're afraid of writing something that's flawed. He advocates free-writing, which involves forcing oneself to write continuously for a period of time and almost inevitably produces useful material [54, p. 3-11].

- **The U.S. Justice System**: Justice in the U.S. has traditionally been based on the idea that defendants are assumed to be innocent until they're proven guilty, with the jury system set up so that they can only be convicted if everyone believes they're guilty. In recent years "innocent until proven guilty" has lost its impact because the dominance of plea bargaining means that people's sentences are based primarily on the strength of the accusations against them rather than the evidence supporting their innocence; this has helped lead to the largest and most racially discriminatory prison system in the world [3, p. 86]. (This example was also noted by Elbow [54, p. 188-9].)

- **Pornography**: Porn star Annie Sprinkle famously said, "The answer to bad porn isn't no porn...it's to try and make better porn!", which is precisely the direction *Good for Her* takes with the annual Feminist Porn Awards. Ending the popularity of porn, especially among people who *aren't* committed feminists, isn't realistic; creating pornography that promotes female empowerment and healthy relationships is. The Feminist Porn Awards recognize pronography that involves "[w]omen and/or traditionally marginalized people", "depicts genuine pleasure, agency and desire for all performers", and "expands the boundaries of sexual representation

on film, challenges stereotypes and presents a vision that sets the content apart from most mainstream pornography" [73].

- **Medicine (and Psychiatry)**: Nassim Nicholas Taleb notes in his book *Antifragile* how harmful medicine has been throughout its history (including present-day) because of it's tendency to over-prescribe[158, p. 341]. This type of interventionism is based on two ideas: that there is something wrong with the patient (which doctors assume any time they want to treat a patient) and that the body is incapable of dealing with it. Both of these are based on critical thinking and problem-seeking. Taleb advocates medical intervention only in extreme cases, where the patient's life (or limb, etc.) is definitely at risk[16]. Psychiatry is similar, both because of the heavy use of actual medications and because of therapy taking the place of medication. Therapy that works with the client (in humanitic psychology), instead of trying to control them or simply stop their behavior, is more effective.

- **Religion**: Unsurprisingly, religion benefits heavily from the belieiving game. While many "intellectuals" and outspoken athetists criticize religion because it can't be proven (and to them the idea of a god is implausible), religious people form communities on that basis, live longer, and are on average happier [113]. Religion has also helped many people with dramatic personal growth, most notably Malcolm X [180]. Regardless of whether God exists, believing in God is helpful to many people.

- **Punishment**: As discussed earlier in this Plan, punishment has been held by most pioneers of behaviorism in psychology to be largely ineffective, largely because it simply tries to negate negative behaviors [93, *Chapter 3*]. Punishment that *is* effective focuses on creating new behaviors to replace old ones (see Guthrie) or in restorative justice focuses on offenders taking responsibility and actively working to repair harm done by their actions.

- **Checklists**: Atul Gawande has noted the recent rise of checklists for dealing with complexity in fields as diverse as construction work, (American) professional football coordination, surgery, flying planes, restaurants, and elsewhere. The essential point about checklists is that they don't try to provide new information to people, who are experts in the field and understand their job at a very deep level. Rather, they act as subtle reminders to people to "ensure that the stupid but critical stuff is not overlooked" [66, p. 79]. By respecting people's expertise and fine-tuned skills, and supporting them in doing basic things they already intend to do (but might forget because of stress and the complexity of the situation), checklists can be enormously useful. It's worth noting at least two major types: content checklists, which focus on the substance of the problem, and communication checklists, which focus on coordinating projects between different people.

- **Climate Change**: Criticisms of climate change have been discussed earlier in this paper, but another important one is that, broadly speaking, the efforts aimed at reducing carbon dioxide both are too expensive to be economically feasible (or desirable) and too small to make a noticeable difference in the climate [107, p. 186-7]. Essentially, the strategy of stopping undesirable and harmful activities is not going work, and reducing carbon dioxide

[16]This essentially corresponds to the idea of "fundamental objection", discussed in terms of Sociocracy and Feminism below.

in the atmosphere might not even be desirable.[17] In terms of addressing climate change, though, the most viable strategies are creating and promoting alternative energy sources and geoengineering. Energy sources could be solar, wind, nuclear, hydropower, or others, while plausible geoengineering solutions include iron fertilization, a stratospheric shield (with sulfur dioxide), and controlled cloud formation. These geoengineering ideas mimic natural processes so that their effects are proven to be harmless.

6.2.3 Statistics and Synthetic Thinking

In statistics synthetic thinking can overcome many barriers, as well as being really cool! I will illustrate this with two examples, one from the Kinsey Reports in 1948 and 1953 and one from a book looking at intelligence in 1972, and then a brief discussion of nonparametric statistics.

6.2.3.1 The Kinsey Reports

In the first example, Alfred Kinsey and a team of researchers were researching human sexuality and conducted what became the largest and most comprehensive study in history (even today), as well as the most influential (in fact, the first sex study of public importance). Kinsey had an incredibly detailed, thoroughly effective methodology, particularly in regard to confidentiality, sampling, and interviewing. Before conducting any of the interviews, Kinsey worked with an "experienced cryptographer" to create a code that would be as difficult to decipher as possible [96, p. 45]. The sheets used to record data had multiple blocks, and both the position of a symbol within the block and which block it's in determine what it means [96, p. 72]. No written record of the code exists, and at no point during the research were the data translated into an uncoded form. Of the six researchers, only 4 of them knew the code in full; the other two each only knew part of the code. In addition, the data were kept in fireproof files behind locked doors, with locks uniquely designed for this study. As a final note, Kinsey talks about the possibility of court subpoenaing the records and says that if the courts insisted on that, "there would be no alternative but to destroy [the] complete body of the record and accept the consequences" [96, p. 47]. Because confidentiality is absolutely vital to this research, Kinsey does everything he possibly can to make sure all the data remain forever anonymous, including creating multiple new security devices (the code and the locks)[18]. As a result, his research team is able to assure subjects that nothing they say can ever be used to incriminate them and that they have nothing to fear about being honest.

Sampling is probably where the Kinsey report encounters its biggest difficulties, primarily because getting truly random samples is one of social science's biggest challenges in general. While statisticians in general usually argue that the possible areas of bias of their sample are unimportant (or fail to address this issue at all), Kinsey used a series of procedures to actively create the most representative sample possible. His sample was the biggest in world history at over 16,000

[17] Ecologist and leading climate scientist Ken Caldeira noted that "[o]verall, more carbon dioxide is probably a *good* thing for the biosphere – its just that it's increasing too fast." [107, p. 185] While this is hardly a ringing endorsement, it points for the need to transition into higher CO2 levels rather than stop them entirely. This is similar to the strategy recommended in *The Transition Handbook* for dealing with climate change and Peak Oil; the author focuses primarily on building resilience and creating a transition plan) [84].

[18] Kinsey notes that "[t]he care with which confidences have been guarded in the present study has probably never been surpassed in any other project dealing with human material." [96, p. 44] While the claim was made in 1948, it's probably still true.

people[19], and there were at least 50 people sampled from each state in the U.S.[20] Noting that truly random sampling would be impossible (both because he had no way of randomly choosing individuals and because some people might not respond), he used a detailed method of stratified sampling that drew from as many different groups as possible. He worked to make the research as diverse as possible with respect to sex, race, marital status, age (ages 3-90 included), education level, occupation, social class, type of city (urban, rural, and mixed), religion, geography, and other factors [96, p. 5][21]. In addition, Kinsey worked to get 100% response rates within each "social unit" of the sample [96, p. 93]. This means that, when sampling people in a college fraternity, Kinsey and his team would work to make sure every single member of the fraternity would be interviewed (because otherwise the sample wouldn't be representative if the people not interviewed differed at all in sexual behavior (this could be because they had extreme sexual histories they didn't want to share, because they were uncomfortable with sex, or any number of other reasons)). At the time of his first study, 26% of the people interviewed were part of these 100% groups, and he includes in the book comparisons of these people to the rest of his sampled population (differences are small, and likely mostly due to the fact that college students are highly overrepresented in the 100 percent groups) [96, p. 96-8]. After sampling, Kinsey did various statistical tests of data validity and found that female inmates as a group had significantly different results from the rest of the population [90, p. 22]. He therefore excluded them from the statistical analyses because they were highly overrepresented, but he used their data in his qualitative discussions, allowing him to get the maximum benefit from the data without negatively affecting the validity of his conclusions[22]. In general, Kinsey took every active step in sampling to make his group as varied and representative as possible, as well as making it large enough that he could draw statistically valid conclusions while analyzing a group in terms of multiple dimensions at once (e.g., comparing single, married, and post-married males in different 5-year age ranges in terms of intercourse with prostitutes [96, p. 288]). This means that his data were nearly as valid as they could possibly be, and became increasingly valid as he continued sampling.

While Kinsey's procedures for ensuring confidentiality and a random sample were incredibly thorough and successful, his interview technique was arguably even better. Because his research required people to give their full sexual histories, there was a definite risk of lying or being unwilling to complete the interview. Due to the nature of interviewing and questions potentially meaning different things to different people, there was also a risk of the data not being standardized. Kinsey solved all of these problems brilliantly by creating an interview structure that was simultaneously

[19]His goal was to keep researching until he got to 100,000; the first book was dedicated to the 12,000 already interviewed and "the eighty-eight thousand more who, someday, will help complete this study." [96, p. *iii*]; by the time of his second book he had data for 16,392 people [90, p. 36].

[20]Alaska and Hawaii weren't states at the time, so this refers only to the 48 contiguous states of the U.S. [96, p. 5].

[21]His "short list of people who have helped most", below, shows some of the diversity of his sample:

"Bootleggers, Clergymen, Clerks, Clinical psychologists, College professors, College students, Corporation officials, Editors, Farmers, Female prostitutes, Gamblers, Headmasters of private schools, Housewives, Lawyers, Male prostitutes, Marriage counselors, Ne'er-do-wells, Persons in the Social Register, Physicians, Pimps, Police court officials, Prison inmates, Prison officials, Professional women, Psychiatrists, Public school teachers, Social workers, Thieves and hold-up men, Y.M.C.A. secretaries, Y.W.C.A. secretaries, Welfare workers, [and] Women's Club leaders" [96, p. 39].

The sample kept growing more diverse with time so that by the end, Kinsey notes that he and his researchers had "a network of connections that could put [them] into almost any group with which [they] wished to work, anywhere in the country" [96, p. 39].

[22]He excluded nonwhites from his statistical analysis also, because he didn't feel he had a large enough sample to make solid concludions (hence the goal of 100,000) [96, p. 6], but used their data for the qualitative analysis as well.

standardized and flexible. The most important measure Kinsey took in this regard was varying the sequence of topics, going from things nonsexual in nature to things the subject would normally find difficult to discuss. Deciding the best order on the basis of general demographic information and individual responses by the subjects, Kinsey and his team were able to ask every subject every question in a way where they would feel comfortable answering. At the time of the first study, 12,000 people had been interviewed and only 3 or 4 did not give a complete history[23]. Kinsey additionally varied the types of words used and the speed of the interview when necessary, particularly for the people interviewed with IQ's between 50 and 70 (and there was also a special process for interviewing children) [96, p. 49; 58]. Additional interview processes, like focusing on objective activities and asking additional questions to cross-check the data, further guaranteed the validity of the interview process, which the entire team was trained in before the study. By creating a totally new interview style (and a personalized interview for every subject), as well as actively building skills in interviewing, Kinsey and his group of researchers were able to get full sex histories from virtually everyone interviewed. Kinsey's synthetic thinking in all parts of the study allowed him to overcome some incredible barriers in statistics and create a study that other researchers have been unable to match in the 60 years since its full publication[24]. The Kinsey Reports became some of the most influential books in U.S. history, directly leading to the Sexual Revolution and the modern LGBT[25] Rights Movement.

6.2.3.2 Ginsburg and *The Myth of the Deprived Child*

My second example comes from Herbert Ginsburg's 1972 The Myth of the Deprived Child. The book in general tries to refute the trend in academia to see poor children as intellectually inferior and promote the idea that they're intellectually capable and should be treated as such, and the example here comes specifically from the section dealing with intelligence testing. At the time, most studies purported to show that poor children were lacking in important intellectual skills (and usually claimed that this was the result of a poor environment or bad parenting). Like most academics, Ginsburg starts with methodological criticisms of some representative studies in the field and questions the validity of the intelligence tests themselves. Ginsburg hypothesizes that the measured class difference in intellectual skills is a result of motivation and comfort with the testing environment, citing research that lower-class children are likely to be less motivated by testing and less familiar with the testing environment [70, p. 38-42]. To that end, he looks at a 1970 study by Frank Palmer that intentionally uses a wide variety of measures[26] to make sure

[23] Additionally, these few refusals were at the beginning of the study; Kinsey claims that if his team had known at the beginning of the study what they knew at the end (in terms of interview technique), even those people would have given full sexual histories. If true, this means that nearly all his interviews (all except the first few hundred) do not suffer from any selection bias from people quitting interviews [96, p. 49].

[24] The next biggest study, *The Hite Report* published in 1981 (most likely only funded because of the Sexual Revolution that followed Kinsey's publication) had only around 15,000 people, and a number of other methodological problems (e.g., a survey was sent to 100,000 females, of which only 4,500 responded; this means that there is a high likelihood of *huge* sample bias [175, p. 172, 180]. Hite was more focused on qualitative things, so the study was still useful, but nowhere close to Kinsey's in terms of statistical validity [82, p. 769].

[25] "Lesbian, Gay, Bisexual, and Transgender". Modern variations are often longer to include more groups; my personal favorite is QUILTBAG: "Queer and Questioning, Unidentified, Intersex, Lesbian, Transgender and Transsexual, Bisexual, Asexual, and Gay and Genderqueer" [173].

[26] This includes taking time to put the child at ease before testing, testing children only as long as they're "alert and responsive", and spending at least nine sessions with each child (despite there being over 300 children in the study), among other things [70, p. 108].

students are fully motivated. He also tests many different skills, "from 'intelligence' as measured by the Stanford-Binet IQ test, to motor abilities" [70, p. 108]. With one exception (the Peabody Picture Vocabulary Test, which favored middle-class children (though it may have been biased in terms of featuring mainly middle-class vocabulary)), there were no class differences. Because the study is large and comprehensive and (unlike other studies) actively takes steps to properly motivate children, its results can be considered more reliable that those of other (individual) studies. By creating a specific procedure to make the data as accurate as possible, Palmer provides strong evidence for a lack of class difference between children in terms of intelligence. Since this study directly addresses Ginsburg's criticism of previous studies, that its results are what Ginsburg predicted provide significant support for his position. So far, so good.

Promising though the Palmer study is, it's ultimately limited by lack of replication, which means that the evidence still overall seems to favor the class difference shown in most other studies. Ginsburg's criticisms of the other studies may be fairly convincing, but his view overall still has little support. After detailing another study with exceptionally strong methodology (that supports his position) and giving his position some theoretical support from Piaget [70, p. 116-129], Ginsburg takes the most important step of the chapter: he decides to use the body of other studies in the field to support his position. This seems impossible since those studies empirical form the basis of the claim that poor children are intellectually inferior, but Ginsburg (after having criticized them) is able to brilliantly show how those studies overall provide significant support to his position. Essentially, he does this by noting that the studies don't show a difference in most areas, meaning they support the claim that (for the most part) there aren't any significant differences between middle-class and poor children [70, p. 131-9]. In many of the cases where there are significant differences, Ginsburg notes that the differences disappear with age, so to the extent that race or class affects cognitive skill, it is in how soon a skill develops only [70, p. 137]. In some of the studies that look at language use, Ginsburg notes that, while differences exist in the frequencies of some "higher-level" linguistic structures, all of the children in each group make use of every structure, meaning the study shows that children of all class backgrounds have these intellectual skills (and this claim is objective and undeniable, at least within this sample). In this analysis Ginsburg epitomizes synthetic thinking in statistics: instead of criticizing and dismissing the results of the studies he disagrees with, he uses as many of those studies as he can to support his claim. With an exceptionally good study from Palmer, theoretical support for his position, and all the data from representative studies in the field supporting his conclusion, Ginsburg singularly makes an extremely compelling case for the idea that poor children are as intellectual capable as other children. Fittingly, this occurs in the context of a larger synthetic thinking-focused book that advocates working with poor children's strengths instead of trying to compensate for their (perceived) weaknesses. Like Kinsey, Ginsburg makes a groundbreaking advancement in the field, almost entirely as a result of statistical analyses motivated by synthetic thinking.

6.2.3.3 Nonparametric Statistics

The Kinsey reports deal with issues of getting good reliable data despite difficulties in meeting assumptions, while Ginsburg's work deals with how to interpret statistics synthetically. One thing neither of these address is how to deal with model uncertainty in statistical inference. As mentioned before, most statistical inference assumes that the relevant population follows a normal distribution, which is rarely completely true and frequently quite untrue. This last method, therefore, is a general model for thinking synthetically in statistical inference and how that problem can be avoided. It

makes use of a branch of statistics called "nonparametric statistics".

'Nonparamtetic statistics' refers to the branch of statistics that does not assume any characteristics about an underlying population. It is therefore the most general form of statistics, essential to inference when the population distribution is unknown (or there aren't statistical methods specific to that distribution type) but accurate in cases where the model is known also.[81, p. 7] Essentially, nonparametric statisticians have developed methods that will work with any type of population. This comes with two main costs: nonparametric statistics are computationally intensive, and their stated statistical significance is lower than that of parametric statistics with the same number of data points.[27] Thes are not huge problems, largely because they allow statisticians to play to their strengths. Computing power has become insignificant to most statistical calculations, and as computing power continues to advance this will become increasingly true [81, p. *xv*]. The reduction in statistical significance can still be a problem, but is tempered by two factors. First, there's the essential point that parametric statistics already have an important level of uncertainty based on how closely the population resembles a normal distribution; in parametric statistics this is hidden, while in nonparametric statistics the actual uncertainty is made explicit. Second, the field of nonparametric statistics is sufficiently developed that it can make use of any information about the population that exists. If the distribution uses actual numbers (instead of just ordinal ranks, like 1st, 2nd, etc.), there is a more powerful nonparametric test available (because the general test can handle even ordinal values). If the data are distributed symmetrically, an even more powerful test can be used to reduce uncertainty further. (In the extreme case, if the data are normal then tests developed for the normal distribution can be used.) There are other information about the data that can be incorporated into the choice of nonparametric test, allowing people to use the strongest statistical test that they are sure fits the data set and then being confident that the uncertainty reported in statistical significance is as low as calculated. Nonparametric tests allow statisticians to be as confident as they possibly can in their results, using information about data set to make conclusions stronger. Nonparametric statistics is synthetic thinking *par excellence* and is directly usable in all areas of statistical inference.

6.2.4 Feminism and Synthetic Thinking

Like statistics, there are some really strong examples of feminism benefitting from synthetic thinking, and the movement as a whole could stand to benefit much more. I discuss racial discrimination and public housing as an example, and then look at sociocracy as an example of how feminism can deal with issues of power hierarchies generally.

6.2.4.1 Housing Discrimination

I would like to propose an alternative to anti-discrimination laws, based loosely on some of the work of anti-racist activist Tim Wise[28] and directly on the centrality of intersectionality to femi-

[27] In some cases, such as basic two-sample inference, the measured statistic significance is essentially the same, and frequently lower with nonparametric statistics! [121, p. 134] In most cases, though, they require larger amounts of data to reach the same level of statistical significance (a good example for someone familiar with the field is the Kolmogorov-Smirnov Omnibus test).[81, p. 57]

[28] In his speech "The Pathology of Privilege", Wise talked about how racism concealed social problems, or more specifically, how the worse economic situation of blacks allowed poor whites to be exploited [178, 35:08]. In Wise's example fighting against racism would help poor whites indirectly by making them harder to exploit; it inspired my idea which helps poor whites by helping poor blacks even more (but without being directly race-focused like Wise's

nism. Anti-discrimination laws are fundamentally taking a critical thinking approach: they view discrimination as the evil and the fundamental goal as preventing it, instead of taking the approach of trying to build a positive structure. Laws against (anti-black) racial discrimination, for example, struggle for multiple reasons, one of which is that they treat blacks as a monolithic group, and act as though the barriers blacks face are primarily (or exclusively) composed of anti-black racism. Intersectionality, the dominant paradigm in feminist theory for the past few decades, asserts that oppression cannot be divided into different characteristics, but that different groups face unique issues. For example, black women face issues that neither black men nor white women face. (An equivalent way of looking at this is that blacks experience sexism differently from whites, or that women experience racism differently from men.) Anti-discrimination laws are unable to address this complexity: if they treat racism as monolithic it won't capture most groups' experiences, while if they treat racism as different for each group the laws will be vague and unenforceable. It's easy to criticize and find cases of discrimination, but it's difficult to make anti-discriminatory public policy. By turning to synthetic thinking, however, it might be possible to create a structure or policy that makes discrimination impossible, or at least relatively unimportant.

To be more concrete, I'll use the specific example of anti-black discrimination with respect to housing. The situation is that blacks are charged substantially more than whites or denied housing, and for the reasons mentioned in the last paragraph anti-discrimination don't seem like they'd be very effective. A synthetic, supportive approach would involve looking at the problem in a different way and trying to build a structure that fulfills blacks' needs. As an example of such a structure, I propose universal housing from the government (which is different from public housing; see footnote.[29]) This effectively eliminates the possibility of the most egregious type of discrimination (denying blacks housing) and severely reduces the importance of the other type of discrimination (higher prices are less important if everyone has a free housing alternative option). In that the program is universal it promotes equality in housing opportunities in a most basic way, while managing to avoid the question of race entirely.[30] This synthetic solution does more than any anti-discrimination law ever could to promote racial equality, without needing to even involve race!

It has even more benefits than that. The previous paragraph established how such a program is more effective than anti-discrimination laws for a given group, but it's also better in that it would apply to more groups. This goes back to intersectionality: blacks might be able to avoid anti-black racial discrimination with an effective anti-discrimination law, but there will still be a large number of blacks who are denied housing (because they're poor, disabled, or otherwise disadvantaged). A universal housing law will help people in all of these groups, meaning blacks as a group would benefit (and many whites would also benefit). All oppressed groups would benefit from this, and

---

idea was).

[29]Public housing, at least as it currently exists, denies large number of groups from using the houses (e.g., felons, drug users) and also severely regulates the conducts of people who live there. In addition, the process of getting public housing can be sufficiently long and difficult that it prevents eligible people from actually receiving the benefits. I think public housing is positive for the reasons described above, but it would be more so if it loosened its requirements to allow more people to use it. (To the extent that it promote segregation, I would respond that this is better than homelessness. Also, improving the quality of public houses would reduce segregation and could benefit even middle class people by making the process of moving substantially easier. Singapore has around 80% of its population in some form of public housing, so having a wide-ranging housing system is not impractical [86].)

[30]This is desirable because it reduces racist backlash and helps other groups as well. In some other circumstance (e.g., academics or testing), avoiding the mention of race while still combating racism is also preferable because it avoids stereotype threat [152, p. 159]

those groups that are marginalized within the black community (e.g. disabled blacks) and that historically haven't benefitted as much from anti-discrimination laws will benefit the most. As a result, this movement could potentially gather much wider support within the black community, and since it would also help many non-blacks (in a way that anti-discrimination laws wouldn't), it could get much more support outside the black community as well. Anti-discrimination laws embody critical thinking and essentially try to end the process of racism by forbidding it. Universal housing (as an example) embodies synthetic thinking and intentionally creates a right and system that benefits more groups inside and outside the black community, and benefits them substantially more than anti-discrimination. Two other examples of such laws come readily to mind: promoting nonviolence instead of creating anti-rape laws (both because rape itself is violent and because most violent rapists were previously victims of major violence), and redefining legal marriage so that it wasn't tied to sexuality or love (which would allow gays to get all the benefits of marriage, as well as close family members who chose to.) Laws such as this, which reduce racism, sexism, or homophobia by creating structures that reduce inequality or social problems in general, are extremely effective because they fundamentally benefit everyone.

6.2.4.2 Sociocracy

Looking more broadly than specific issues, feminism is deeply concerned with issues of power and representation, and therefore governance. Given the centuries-long oppression of women, blacks, the poor, the disabled, and other marginalized groups, and the fact that they continue to be shut out from roles of power today, our current system can in some ways be considered a dictatorship (and indeed, much anti-sexist and anti-racist rhetoric is to that effect). In this sense feminists should be using synthetic thinking to change our system of government already, given its centrality to overcoming a dictatorship in Sharp. More specifically, though, synthetic thinking can be used both to create and to operate a general system of governance: a relatively new system called "sociocracy" illustrates this well. "[A]lso known as dynamic governance or dynamic self-governance" [28, p. 7], sociocracy is a consent-based government model whose main operating feature include double-linked governance circles, consent-checking rounds, and specific processes for elections.[31] I will describe

[31]One of the other major features described in *We The People* is monetary feedback, or sort of a combination financial incentive system and feedback system. This is in my judgment less essential than the other features (in particular, I think sociocracy could work in volunteer or other organizations without pay), and its merits are more economic than socio-political. But for interested readers, here is how it works:

Essentially, everyone has a base salary but are able to make more than that if they contribute to the company [28, p. 171]. Whenever the company makes more than projected, an elected financial officer figures out which employees extra work or innovations caused this excess and divides the money between them according to their impact. If someone underperforms during a period, they still receive their base pay but the amount of lost income is taken from their future "bonuses" until the debt is paid. The idea behind this system is that people receive objective (monetary) feedback about the impact of their changes in their behavior and how their work has contributed to the company [28, p. 167]. They are pushed to work harder because it directly benefits them financially; if money is less important to them they may choose not to put extra time into work, and that's fine (as long as their work is adequate) because it doesn't affect anyone else's money.

This system is difficult to implement in part because assessing the impact of an individual is difficult to do reliably. A company can default to dividing any extra money among all employees equally (or any who participated in a new project, etc.) if it chooses to do this, and the financial officer in charge of allocating surplus profits could be forced to quit their position if it was felt they were biased (or just wrong). It's also worth noting that employees could choose to invest extra profits in the company, collectively or individually; this would ideally reflect in extra money they earn in future pay periods [28, p. 172].

each of these features[32], along with how they embody synthetic thinking and how they meet the goals of feminists.

At its most basic level, sociocracy is based on the idea of consent at all levels of decision-making. Consent here means "the absence of paramount objections"; in other words, the acknowledgement that decision does not create problems that cause people to not function [28, p. 43; 69]. This is different from disagreement, in that a person may disagree with a decision without having a paramount objection. For example, I might object to my organization taking time to celebrate people's birthdays because I'd rather be working (and I hate birthday celebrations, my own or others). However, unless my work requires me being on call during that time or something similar[33], this is not a paramount objection, meaning I "consent" to the decision in sociocratic terms (but still don't like it!). All decisions are made according to consent in a sociocracy, even elections. The election procedure essentially involves people nominating themselves or other for positions (and giving a reason for their nomination), giving other a chance to opine about different candidates, and then having the circle facilitator propose a candidate and seek out paramount objections (the process is described in more detail in this footnote[34]). As with all decisions, the facilitator (who is also elected) actively seeks out paramount objections, prompting people if they seem uncomfortable with a decision even if they've nominally consented. As a process for decision-making this means that everyone is represented at least in regards to paramount objections, so that no one's needs can be ignored. This is enormously helpful for combating oppression, and like systems from the Tim Wise discussion above, it benefits all types of women or other groups (instead of just middle-class, white, able-bodied women). It's also extremely synthetic: it specifically does not allow criticisms of a proposed policy to change whether it's adopted unless they are of the utmost importance [28, p. 67]. (These criticisms are taken into account in formulating the policy, however.) Philosophically people are assumed to implicitly support any decision reached by a circle, so that support comes from the entire community [28, p. 140-1]. Theoretically and practically, the sociocratic understanding of consent is based on synthetic thinking and avoids critical thinking except when absolutely necessary or helpful. It also represents all of people's most important needs and therefore prevents oppression from occurring in the system.

Just as consent is the philosophical basis for sociocracy, circles are the technical basis. At every level of an organization or government, people are part of a decision-making circle. Each circle has an operations leader (who supervises the activity of people in the circle and is appointed by

[32]At the same time, this model of governance is merely meant to be illustrative; I advocate feminists adopting a different model if they find one that works better. (I do think sociocracy has a lot of potential, though, in part because it's flexible and can work for businesses, organizations, and general governance. I would certainly advocate using it until they create a better system.)

[33]Like doctor Daniel Kalla [92, *inside back cover*]!

[34]First, everyone is given a piece of paper for them to write their name and who they're nominating (everyone is expected to nominate someone). They write these and pass them in. Once all the papers are received, the facilitator reads them one at a time, asking the person to give an explanation for why they nominated the person they did. After that, there is at least one round in the circle in which people may raise objections or decide to change their nominations. When they feel it's appropriate, the facilitator will then propose someone to be appointed to the position. There is another objections round for people to make paramount objections. Once those are resolved (or another candidate is found for whom there are no paramount objections), the next position is elected (if applicable) [28, p. 150-4]. As a note, it's important in this process that the facilitator not ask people ahead of time who's interested in the position; it should be decided by the circle, rather than based on who is interested. Interested parties can nominate themselves if they can justify it, but the election may go to someone else who does not expect to get nominated originally (or might expect it and be willing to accept, but would not volunteer for the position or nominate themselves) [28, p. 153].

the circle immediately higher in the hierarchy), a facilitator, and representatives from and to the circles immediately above and below them in the hierarchy [28, p. 128]. In a circle meeting, the facilitator starts a round by saying what it's going to be about and then choosing a person to start [28, p. 136]. This person says something and then the person next to them has a chance to speak, repeated until everyone in the circle has spoken (at which point the facilitator starts another round, asks someone with a proposal to present that, or ends the meeting). Everyone is given a chance to speak, although people may pass if they choose to do so. A common format for circle meetings is an opening round, a proposal being presented (and followed by clarifying questions), a reaction round, a consent round, and a closing round (or more proposals) [28, p. 135-142]. The two representatives speak to the concerns of people of higher and lower circles so that decisions are more integrated in the organization, and two other representatives elected by the circle convey their concerns to the circles immediately above and below (the system is double-linked (sending a representative to a circle and receiving a different representative from that circle) to make sure that each circle is adequately represented and to avoid possible conflicts of interest) [28, p. 86-7]. In the circle process, everyone is represented and able to give their opinion, even people from the other circles. The system is set up so that everyone by default will contribute and be heard, unless they intentionally decide not to (by deciding to pass). Once something is said it is considered to belong to the group rather than the individual who said it, and the attitude is solution-seeking rather than problem-seeking.[35] This means that whatever people say will be listened to and seriously considered, and they won't be criticized or ostracized for saying something that goes against the group. By encouraging everyone to speak and contribute, and listening to them when they do, the system automatically works to promote everyone's input. Within the circle there is an equality of power and genuine support extending to every member. For feminism and similar movements to end oppression and empower the historically disempowered, sociocracy or a similar synthetic thinking-oriented model is extremely valuable.

I have argued in this section that synthetic thinking is extremely valuable, both because it's essential in many fields and because it outperforms critical thinking alone. Synthetic thinking is the basis for conflict resolution, which is an essential skill to all parts of life, and also essential for overthrowing a dictatorship, which is vital to peoples of a decreasing but still far too large number of countries. In statistics, synthetic thinking provides the best inroads against the problems of certainty that can never be perfectly solved and provides an unmatched strength in statistical conclusions, in place of or alongside critical thinking. In feminism, synthetic thinking is uniquely necessary to build the structures that promote equality and empowerment and can build support for a system over all but the most important of objections. These issues of conflict, knowledge, and power affect all aspects of life and show the importance of synthetic thinking directly, as well as serving as examples of what synthetic thinking can accomplish generally. To end this paper, I detail a synthetic thinking model of deciding between two opposing truths, intended to replace the more common critical thinking model of testing both ideas for flaws. In doing so, I hope to help readers convinced of the superiority of synthetic thinking to be able to use it to their full advantage.

[35] Even when people give objections in consent rounds, the attitude is solution-seeking rather than problem-seeking: "[i]n consent rounds, objections are ot only welcome but sought out. They are not considered barriers to moving forward but opportunities to resolve problems before they obstruct the execution of policy." [28, p. 140] It is important that synthetic thinking in general does not try to silence objections or criticism, but rather to use them productively and find solutions.

6.3 Deciding between Two Claims with Synthetic Thinking

The basic method of deciding between two contradictory claims has already been alluded to in the conflict resolution discussion. Essentially, the technique involves believing fully in both sides, until either someone can reconcile them and decide that they're both right, or one is preferred because there is much more overall support for it. I will briefly describe each approach theoretically and then illustrate it with an example, starting with the first approach.

6.3.1 MRA and gender roles

As I mentioned earlier in discussing the importance of synthetic thinking, believing the other person's argument is essential to understanding it. This does not mean believing it inflexibly and dogmatically, because anyone searching for truth must be willing to sometimes discard knowledge that they believe to be true. Quite often, however, apparent contradictions are not inherent contradictions[36], as seemingly opposite views both describe part of the truth and can be reconciled for better understanding overall. Management guru Peter Drucker agrees, saying that you should "assume that each faction gives the right answer, but to a different question...Then, you gain understanding. You also gain, in many cases, the ability to bring the two together in a synthesis." [51, p. 124] I would argue that the existence of a contradicting theory is positive, because if both are supported by evidence it means that either can be improved with understanding from the other. To illustrate this I use an example from Men's Rights Activists, a group known for claiming that feminism is harmful and that men face high levels of discrimination. Because most of the audiences I write for are (substantially) more sympathetic to feminism than MRAs and believe that women are oppressed historically and currently, and also because I'm a feminist, I write from the perspective of trying to understand the MRA argument while already believing feminist doctrine (however, the argument could be remade the other way quite easily). (For extra fun, and because the argument is related, I also cite self-described conservative feminist Christina Sommers.)

To support their contention that men rather than women are oppressed in contemporary society, MRA groups point to discrimination against males in adoption, child custody, insurance, the military, and other institutions. They also point to how males are more often imprisoned, and more likely to drop out of school [150, p. 32-3]. They blame this difference on the feminist movement, claiming that it has resulted in discrimination against males. Since females legally have the same rights as males, and since males fair worse in these areas, our culture is oppressive toward males. MRA groups often additionally claim that feminism is harmful and disempowering to females, who are pushed from comfortable female roles into male roles against their will. This is supported mainly by anecdotal evidence, although they are usually able to find large amounts of this evidence from their own and their friends' lives. Obviously on the surface this contradicts feminism, both in terms of denying females are oppressed and in terms of saying feminism is negative. Under the synthetic thinking model, someone trying to reconcile these two views tries to believe both of them at once and see how aspects of both of them are true. Fortunately, it's fairly easy to believe

[36]In principled negotiation this is an essential part of Fisher and Ury's framework, and it is the basis for Stephen Covey's The 3rd Alternative [40, p. 9], which also deals with conflict resolution. In Inside the Box, Boyd and Goldenberg talk about how apparent contradiction is essential, or at least extremely helpful, for creativity:

"spotting a contradictionis a very exciting moment, because it fuels enormous creativity: contradiction is a blessing. It is a pathway to creativity...behind every contradiction is an untrodden path that leads directly to options and opportunities that may not have been considered" [25, p. 191].

the MRA argument, because they have a lot of evidence and well-developed explanations of their positions.

Once the MRA position is understood (and the feminist position is also), it's not particularly hard to reconcile them, in part because many people already have. Broadly speaking, we can see that the discrimination males face in these areas *is the direct result of gender roles and stereotypes*: females are assumed to be more involved parent figures and therefore favored in adoption and custody cases, while males are assumed to be more violent and therefore have higher insurance premiums, are drafted into the military, and are more likely to go to prison. Since females also face discrimination is various areas as a result of gender roles and stereotypes, the two groups can be allies and work against gender roles together. Both groups benefit from this arrangement, as both have their grievances addressed. Note that this is true regardless of whether females actually are better at childcare or males actually are more violent: if these difference don't exist, it's due to stereotypes and discrimination, while if they do exist (and it would be surprising if they didn't exist statistically, given the socialization of gender roles), the differences in treatment are valid. Creating more equal gender roles will reduce both perceived and actual differences, benefitting both groups.

The other MRA claim, that feminism is disempowering to women and bad for society, can also be easily reconciled with the idea that it's fundamentally good and empowering to women. In fact, that already happened earlier in this paper! The goals of feminism are extremely empowering, but as we've seen, the methods it uses (in particular, its basis in critical thinking) mean it can be disempowering for many people. I have multiple female friends who are feminists and have expressed guilt about wearing makeup or performing stereotypically feminine acts, and many blog posts say something similar about the pressures individual female authors feel to not be stay-at-home mothers [76, 104, 123]. By viewing feminist goals as empowering but their methods as (sometimes) disempowering, the two groups can agree about the feminist movement and work together to create a more consistently empowering path to social equality. Again, the differences between the two sides lead to a greater understanding of the problem and suggest a way for the two groups to agree and work together. Other examples abound, and this model can be used to resolve disagreements in a wide variety of circumstances.

6.3.2 Picking a Paradigm in the Natural Sciences

The other way to decide between two different paradigms is to see which is best supported, either theoretically or in terms of usefulness. George E. P. Box once said, "All models are wrong, but some are useful," and this idea is an integral part of systems thinking. Any paradigm will be wrong: if not inaccurate it's at least incomplete, and usually both. In the case of extreme, irreparable flaws (the intellectual equivalent of "paramount objections"), critical thinking can show that one paradigm is wrong, but generally paradigms are decided by how well they work, how theoretically or practically "useful" they are as models. It's here, though, that critical thinking is probably most often useful. While it doesn't work by itself to choose one model over another or to discard one in most circumstances, critical thinking does work to show what some models fail to do, and therefore what another model should try to do. Critical thinking can show how statistics can be biased so Kinsey knows to create a unique interview style and Ginsburg knows that the Palmer experiment is more reliable than the other tests; it shows how blacks are disproportionately denied housing so they know to organize or how democracy fails to represent everyone so the need for sociocracy is clear. Critical thinking, in short, can direct attention so that people know a change is needed, as well as what to focus on when evaluating different paradigms. In this way, by motivating and

directing synthetic thinking so it can be most effective, critical thinking can do the most good and make full use of the development it's undergone from being a focus of academia for so long. My final example goes back to pure natural science with the scientific method that forms the basis for critical thinking, because it is here that critical thinking and synthetic thinking have worked harmoniously and here that Western civilization has been able to truly advance, more than any other area. To (finally!) end this paper, I unite critical and synthetic thinking through the lens of Thomas S. Kuhn's *The Structure of Scientific Revolutions* and the history of Western science.

After a brief discussion about history, Kuhn starts his seminal text by describing the practice of 'normal science', or "research firmly based upon one or more past scientific achievements, achievements that some scientific community acknowledges for a time as supplying the foundation for its further practice" [102, p. 10]. As Kuhn's ultimate goal is describing scientific revolutions, normal science' can also be understood as science practiced in any period except those during (or immediately preceding) a scientific revolution. In normal science, Kuhn articulates three main focuses of scientific investigation. First, the paradigm is used to explain and predict natural phenomena (e.g. the position of planets or the timing of an eclipse); this is what gives the paradigm its real purpose and what makes it useful outside of academia [103, p. 25]. Second, investigations are made to test (or more often, to confirm) the paradigm; a good example of this is Foucault's pendulum.[37] [103, p. 26] Third, investigations are made to extend or clarify the paradigm; for example, scientists may design a larger-scale experiment to measure the gravitational constant G more precisely than ever before [103, p. 27-8]. Essentially, the first focus increases the practical usefulness, the second provides theoretical support, and the third extends the model. In normal science these three focuses have the effect of further entrenching the model, while also making it much more developed and useful. Through the process of conducting normal science, scientists have made enormous progress in all areas of natural science, especially in terms of developing technologies. In this way, normal science accounts for the vast majority of scientific progress.

In the heart of his book, Kuhn describes the process by which normal science undergoes a revolution. Contrary to scientific philosophy, contradictory evidence is not sufficient to cause scientists to abandon a paradigm: instead, either contradictory evidence is not noticed, it is considered a puzzle for future scientists, or the theory is clumsily patched to incorporate it[38]. Indeed, in Kuhn's process scientists only begin to search for alternate paradigms when a very large amount of evidence contradicts the paradigm [103, p. 67-8]. Even then, the paradigm will not be discarded until a different paradigm can replace it [103, p. 77]. This new paradigm must fix the old problems while creating fewer or smaller ones of its own, and it must be consistent with existing technologies that have come out of the other paradigm. When this happens, the paradigm shifts and normal science can resume, now with a bunch of new issues to explore, questions to answer, and a paradigm to articulate. In this way a revolution serves two purposes: it creates a paradigm that describes facts better than (or at least as well as) the previous paradigm, and it allows normal science to continue

[37] "Foucault's pendulum" refers to an experiment in which a tall pendulum that was able to swing on any vertical plane starts swinging, with marks on the ground below it in a circle that show the direction it started and other directions. Because of the earth's rotation, after 2 hours the pendulum bob will be offset from its starting point by about 30 degrees [151]. Since this was created once people already accepted that the earth rotated, this served more as a way of showing that conveniently and elegantly than actual theory building. It provided evidence for the contemporary scientific paradigm, although it was intended as an illustration/confirmation rather than a test of the paradigm at the time.

[38] A good example of the first of these is when electrons were recorded as "bouncing off" electrically charged surfaces, even though current scientific understanding sees this as repulsion [103, p. 117]; an example of the third one is the Greek invention of retrograde motion to explain why planets and the sun didn't orbit the earth in perfect circles.

by overcoming the major flaws of the previous paradigm and generating a variety of new directions to explore. With a successful scientific revolution, science not only redeems itself but also maintains its ability to continue to advance.

The continued development of science, one of humanity's greatest technological and intellectual achievements, is successful because it uses synthetic thinking in a masterful way and supports this with critical thinking only as appropriate. Despite the history of scientific revolutions (which Kuhn notes are much more common than scientists tell themselves [103, p. 6-8]) and the consequent high likelihood that the current model of science is simply false[39], scientists accept and believe that their current paradigm is correct and practice normal science. Here, ignoring contradictory data[40] and believing in the paradigm leads to the vast majority of scientific progress, while accepting the criticisms would actively inhibit the ability of science to advance.[41] When the number of contradictions become too large and begin to prevent scientists from being able to practice normal science effectively, they act like a paramount objection (Kuhn uses the term "crisis") and *force* scientists to look for a solution. However, synthetic thinking still remains the most important, because scientists do not discard a paradigm until they can replace it with a better one. Revolutions work to create new paradigms that both allow science to keep its technological and theoretical advances, while also providing a space to move forward. This means, in effect, that science as a whole cannot ever move 'backwards' and must always move forwards. While a large part of the success of Western science is due to its focus on objective phenomenon (on which it's usually possible to conduct experiments), it is essential to note how the focus on synthetic thinking both in practicing normal science and in developing new paradigms allows science to continue to advance. It remains a definitely flawed and wrong institution, but by advancing synthetically and only using critical thinking to guide a revolution when one is absolutely necessary, it has made incredible advances in human understanding. At the same time, the critical thinking inherent in the falsifiability requirement of science prevents worse, harmful theories from becoming a part of the scientific paradigm.

Thus, synthetic thinking accounts for the development of science within a paradigm and the creation of between paradigms during a revolution, while critical thinking shows when a revolution is necessary and protects the progress of science. In this case, the conservative nature of critical thinking is a bonus, because the synthetic thinking in science means that the field is able to keep progressing. Thus, synthetic thinking can lead to progress while critical thinking can give a direction to that progress and protect progress from degrading. Together, these types of thinking have supported the incredible progress of Western science, and by using the same model they can do similarly revolutionize other areas of scholarship or social change. This model of thinking has

[39]Not only that; because science relies on experimentation and falsifiability, it has to be possible for the model to be wrong. If it were certain, it wouldn't be "scientific" in a strict sense. So this is literally unavoidable for any theory in the current model of science.

[40]For a current example of this phenomenon, consider dark matter. Dark matter is essentially matter that's hypothesized to exist but visible only though it's gravitational effects, which are substantial. Scientists estimate that around 85% of the matter in the universe in dark matter, which is possibly uniformly distributed (but it's unknown since it has no noticeable effects other than gravity) [140]. You could look at it as an extension of the model that makes it more useful, and many scientists do. My personal tendency is to look at it as the model under-predicting the effect of gravity by a factor or seven for certain large cosmic interaction (e.g. they predict it's 1/7 as strong as it actually is), and then creating a post-hoc explanation so it's not discarded. While the case of dark matter specifically is unknown, there certainly are cases where science models are adjusted like this to prevent evidence from contradicting the model.

[41]Kuhn quotes Francis Bacon as saying that "[t]ruth emerges more readily from error than from confusion." [103, p. 19]

enormous implications for every discipline, and a lot to contribute to their development.

In conclusion, a strong primary focus on critical thinking is extremely harmful and limiting, as illustrated by examples from statistics and feminism. Synthetic thinking is essential for progress in many aspects of life and much more powerful alone than critical thinking, in areas as diverse as conflict resolution, nonviolent revolution, statistics, and feminist critical theory. Synthetic thinking also serves to best help someone decide between (or among) conflicting viewpoints and use disagreements to advance their own knowledge. Synthetic thinking appropriately supplemented by critical thinking is even more powerful than synthetic thinking alone, however. This can be seen in statistics and in the feminist movement, and it forms the basis for the entire Western science paradigm. I have argued that synthetic thinking must be promoted, not only alongside critical thinking but actually to the point of replacing it as the primary type of thinking emphasized in schools. Structures for using synthetic thinking are at least as useful as models of critical thinking, and the skills are overall much more essential. I hope the reader has enjoyed this discussion and can use these models of synthetic thinking to better their own understanding of the world and their own lives.

Bibliography

[1] C(lie)mate #3 - what doesn't make the news, November 2013 (accessed March 28, 2014). URL http://www.youtube.com/watch?v=2hOhq0vHfMI&feature=youtube_gdata_player.

[2] C(lie)mate #2 - CO2 & solar connection, November 2013 (accessed March 29, 2014). URL http://www.youtube.com/watch?v=Hob3-XCxdEk&feature=youtube_gdata_player.

[3] Michelle Alexander. *The new Jim Crow: mass incarceration in the age of colorblindness*. New Press ; Distributed by Perseus Distribution, New York; [Jackson, Tenn.], 2010. ISBN 9781595581037 1595581030 9781595586438 1595586431.

[4] Saul David Alinsky. *Rules for radicals: a practical primer for realistic radicals*. Vintage Books, New York, 1989. ISBN 0679721134 9780679721130.

[5] Ashley Alman. A columbia professor reveals the shocking reality of drug addiction, March 2014 (accessed April 9, 2014). URL http://www.huffingtonpost.com/2014/03/06/drug-addiction_n_4915197.html?ncid=fcbklnkushpmg00000021&ir=Crime.

[6] Eric Alterman. The conservative war on knowledge, May 2012 (accessed April 21, 2014). URL http://www.americanprogress.org/issues/media/news/2012/05/24/11566/think-again-the-conservative-war-on-knowledge/.

[7] Amy Babineaux-Smith. Does ISS (in school suspension) really work? students sound off, May 2010 (accessed April 21, 2014). URL http://www.examiner.com/article/does-iss-school-suspension-really-work-students-sound-off.

[8] Gloria Anzaldua. *Borderlands = La frontera*. Aunt Lute Books, San Francisco, 1999. ISBN 1879960575 9781879960572 1879960567 9781879960565.

[9] Association of College & Research Libraries. Information literacy competency standards for higher education, 2014. URL http://www.ala.org/acrl/standards/informationliteracycompetency.

[10] Lonnie H Athens. *The creation of dangerous violent criminals*. University of Illinois Press, Urbana, 1992. ISBN 0252062620 9780252062629 0252019393 9780252019395.

[11] Lonnie H Athens. *Violent criminal acts and actors revisited*. University of Illinois Press, Urbana, 1997. ISBN 0252023064 9780252023064 0252066081 9780252066085.

[12] Abhijit Banerjee. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. PublicAffairs, 2012. ISBN 9781610390934 1610390938.

[13] Willie Baptist and Jan Rehmann. *Pedagogy of the poor: building the movement to end poverty*. Teachers College Press, New York, 2011. ISBN 9780807752289 0807752282 9780807752296 0807752290.

[14] Cesare Beccaria. *Of Crimes and Punishments*. Blackmask Online, 2001 (accessed April 28, 2014). URL `http://criminologytoday.com/beccaria.pdf`.

[15] Cesare Beccaria and Voltaire. *An essay on crimes and punishments*. Printed for J. Almon, London, 1767 (accessed April 28, 2014).

[16] Glenn Beck, Kevin Balfe, and Steve Burguiere. *An inconvenient book: real solutions to the world's biggest problems*. Threshold Editions, New York, 2007. ISBN 9781416552192 1416552197.

[17] Jennifer E Beer, Eileen Stief, and Friends Conflict Resolution Programs. *The mediator's handbook*. New Society Publishers, Gabriola island, BC, 1997. ISBN 0865713596 9780865713598 0865713588 9780865713581.

[18] Jillian Berman. What obama ignored about the 'Lowest unemployment rate in over 5 years', January 2014 (accessed April 27, 2014). URL `http://www.huffingtonpost.com/2014/01/28/obama-unemployment-rate_n_4684396.html`.

[19] Neil Blomberg. Effective discipline for misbehavior: In-school versus out-of-school suspension*. *CONCEPT*, 27(0), November 2003 (accessed April 21, 2014). URL `http://concept.journals.villanova.edu/article/view/138`.

[20] Chris Boccanfuso and Megan Kuhfeld. Alternatives to zero tolerance policies, 2011 (accessed April 21, 2014). URL `http://www.nea.org/home/alt-zero-tolerance-policies.html`.

[21] Erling E. Boe and Russell M. Church. *Punishment; Issues and Experiments*. Ardent Media, 1968.

[22] R Bolton. *People skills: how to assert yourself, listen to others, and resolve conflicts*. Simon & Schuster, New York, 1979. ISBN 067162248X 9780671622480.

[23] Wayne C Booth, Gregory G Colomb, and Joseph M Williams. *The craft of research*. University of Chicago Press, Chicago, 1995. ISBN 0226065839 9780226065830 0226065847 9780226065847.

[24] Cornell Dewey-Jimerson Shane Borum, Randy. What can be done about school shootings? a review of the evidence. *Educational Researcher*, 39(1):27–37, 2010. ISSN 0013-189X.

[25] Drew Boyd and Jacob Goldenberg. *Inside the box: a proven system of creativity for breakthrough results*. 2013. ISBN 9781451659252 1451659253.

[26] John Braithwaite. Shame and criminal justice. *Canadian Journal of Criminology*, 42(3):281–298, July 2000 (accessed November 15, 2013). ISSN 07049722. URL `http://libproxy.marlboro.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=3412728&site=ehost-live`.

[27] William J Bratton, Norman Dennis, Institute of Economic Affairs (Great Britain), and Health and Welfare Unit. *Zero tolerance: policing a free society*. IEA Health and Welfare Unit, London, 1997. ISBN 0255364326 9780255364324.

[28] John A Buck and Sharon Villines. *We the people: consenting to a deeper democracy : a guide to sociocratic principles and methods*. Sociocracy.info, Washington, DC, 2007. ISBN 0979282705 9780979282706.

[29] Bureau of Labor Statistics. Earnings and unemployment rates by educational attainment, March 2014 (accessed April 27, 2014). URL `http://www.bls.gov/emp/ep_chart_001.htm`.

[30] Alberto Cairo. *The functional art: an introduction to information graphics and visualization*. 2013. ISBN 9780321834737 0321834739.

[31] Christina Caron and Gillian Mohney. Va. tech response: Better than 2007, December 2011 (accessed April 25, 2014). URL `http://abcnews.go.com/US/virginia-tech-improves-security-response-2007-shootings/story?id=15114721`.

[32] Sarah Carr. Do 'Zero tolerance' school discipline policies go too far? *Time*, May 2012 (accessed April 13, 2014). ISSN 0040-718X. URL `http://content.time.com/time/nation/article/0,8599,2115402,00.html`.

[33] A. Charles Catania. *Contemporary research in operant behavior*. Scott, Foresman, Glenview, IL., 1968.

[34] Erica Chenoweth and Maria J Stephan. *Why civil resistance works the strategic logic of nonviolent conflict*. Columbia University Press, New York, 2011. ISBN 0231527489 9780231527484.

[35] CNN. Analysis: Jobs matter, unemployment rate doesn't, May 2011 (accessed April 27, 2014). URL `http://news.blogs.cnn.com/2011/05/06/analysis-jobs-matter-unemployment-rate-doesnt/`.

[36] Wilson M. Compton, Deborah Dawson, Sarah Q. Duffy, and Bridget F. Grant. The effect of inmate populations on estimates of DSM-IV alcohol and drug use disorders in the united states. *The American journal of psychiatry*, 167(4):473–474, April 2010 (accessed March 28, 2014). ISSN 0002-953X. doi: 10.1176/appi.ajp.2009.09081087. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917586/`. PMID: 20360330 PMCID: PMC2917586.

[37] J E Conklin. *Criminology*. United States, 1995. ISBN 0-02-323811-9.

[38] Lewis A Coser. *The functions of social conflict*. Free Press, Glencoe, Ill., 1956.

[39] Stephen R Covey. *The 7 habits of highly effective people: restoring the character ethic*. Free Press, New York, 2004. ISBN 9780743269513 0743269519 9780743272452 0743272455 9781417656646 1417656646.

[40] Stephen R Covey and Breck England. *The 3rd alternative: solving life's most difficult problems*. Free Press, New York, 2011. ISBN 1451626274 9781451626278 9781451626261 1451626266 9781451626285 1451626282.

[41] Jeff Cox. Why the unemployment rate has become a bad joke, February 2011 (accessed April 27, 2014). URL http://www.cnbc.com/id/41583533.

[42] Crime and Punishment in Durham, 1750-1900. The bloody code, 2014 (accessed April 28, 2014). URL http://community.dur.ac.uk/4schools/Crime/Bloodycode.htm.

[43] Anne E. Cunningham and Keith E. Stanovich. What reading does for the mind, 2001 (accessed May 5, 2014). URL http://www.csun.edu/~krowlands/Content/Academic_Resources/Reading/Useful%20Articles/Cunningham-What%20Reading%20Does%20for%20the%20Mind.pdf.

[44] Angela Y Davis. Masked racism: Reflections on the prison industrial complex, September 1998 (accessed April 13, 2014). URL http://colorlines.com/archives/1998/09/masked_racism_reflections_on_the_prison_industrial_complex.html.

[45] Angela Y Davis. *Are prisons obsolete?* Seven Stories Press, New York, 2003. ISBN 1583225811 9781583225813.

[46] Michael L. Davis. Time and punishment: An intertemporal model of crime. *Journal of Political Economy*, 96(2):383–390, April 1988 (accessed April 13, 2014). URL http://www.jstor.org/stable/1833115.

[47] Edward L Deci and Richard M Ryan. *Intrinsic motivation and self-determination in human behavior*. Plenum, New York, 1985. ISBN 0306420228 9780306420221.

[48] Barbara Deming. *Prison notes.* Grossman Publishers, New York, 1966.

[49] dictionary.com. the definition of feminism, 2014. URL http://dictionary.reference.com/browse/feminism.

[50] Cathie Douglas and Larry Moore. The healing river an invitation to explore restorative justice values and principles., 2004.

[51] Peter F Drucker. *Managing the non-profit organization.* Butterworth-Heinemann, 1990. ISBN 0750600179 9780750600170.

[52] Emily Dufton. The war on drugs: How president nixon tied addiction to crime, March 2012 (accessed April 20, 2014). URL http://www.theatlantic.com/health/archive/2012/03/the-war-on-drugs-how-president-nixon-tied-addiction-to-crime/254319/.

[53] Emile Durkheim. *Moral education; a study in the theory and application of the sociology of education.* Free Press, New York, 1973.

[54] Peter Elbow. *Writing without teachers.* Oxford University Press, New York, 1998. ISBN 0195120167 9780195120165.

[55] Peter Ferrara. Obama's real unemployment rate is 14.7%, and a recession's on the way, October 2012 (accessed April 27, 2014). URL http://www.forbes.com/sites/peterferrara/2012/10/11/obamas-real-unemployment-rate-is-14-7-and-a-recessions-on-the-way/.

[56] Ford Fessenden. Florida list for purge of voters proves flawed. *The New York Times*, July 2004 (accessed April 8, 2014). ISSN 0362-4331. URL http://www.nytimes.com/2004/07/10/us/florida-list-for-purge-of-voters-proves-flawed.html.

[57] Lydia Alix Fillingham and Moshe Susser. *Foucault for beginners*. Writers and Readers Pub., New York, 1993. ISBN 086316160X 9780863161605.

[58] Roger Fisher, William Ury, and Bruce Patton. *Getting to yes: negotiating agreement without giving in*. Penguin Books, New York, N.Y., 1991. ISBN 0140157352 9780140157352.

[59] Roger Fisher, William Ury, and Bruce Patton. *Getting to yes: negotiating agreement without giving in*. Penguin, New York, 2011. ISBN 9780143118756 0143118757.

[60] Michel Foucault. *Discipline and punish: the birth of the prison*. Vintage Books, New York, 1979. ISBN 0394727673 9780394727677.

[61] Kathryn J. Fox. Circles of support & accountability: Final report prepared for the state of vermont department of corrections. Technical report, University of Vermont, November 2013 (accessed April 28, 2014). URL http://www.doc.state.vt.us/about/reports/circles-of-support-accountability-final-report/view/end.

[62] Paulo Freire. *Pedagogy of the oppressed*. Continuum, New York, 2000. ISBN 0826412769 9780826412768.

[63] P. S. Fry. *Psychological Perspectives of Helplessness and Control in the Elderly*. Elsevier, December 1988. ISBN 9780080867113.

[64] David Garland. *Punishment and modern society: a study in social theory*. University of Chicago Press, Chicago, 1993. ISBN 0226283828 9780226283821.

[65] David Garland. *The culture of control: crime and social order in contemporary society*. University of Chicago Press, Chicago, 2001. ISBN 0226283836 9780226283838 0226283844 9780226283845.

[66] Atul Gawande. *The checklist manifesto: how to get things right*. Picador, New York, 2010. ISBN 9780312430009 : 0312430000.

[67] Generalitat de Catalunya. El franquismo en cataluña, 2010 (accessed April 28, 2014). URL http://www20.gencat.cat/portal/site/culturacatalana/menuitem.be2bc4cc4c5aec88f94a9710b0c0e1a0/?vgnextoid=841c5c43da896210VgnVCM1000000b0c1e0aRCRD&vgnextchannel=841c5c43da896210VgnVCM1000000b0c1e0aRCRD&vgnextfmt=detall2&contentid=0e613c084ded7210VgnVCM1000008d0c1e0aRCRD&newLang=es_ES.

[68] James Gilligan. *Violence: reflections on a national epidemic*. Vintage Books, New York, 1997. ISBN 0679779124 9780679779124.

[69] John Gilliom and Torin Monahan. *SuperVision: an introduction to the surveillance society*. The University of Chicago Press, Chicago, 2013. ISBN 9780226924434 0226924432 9780226924441 0226924440.

[70] Herbert Ginsburg. *The myth of the deprived child: poor children's intellect and education.* Prentice-Hall, Englewood Cliffs, N.J., 1972. ISBN 0136091563 9780136091561 0136091490 9780136091493.

[71] Henry A Giroux. *Stormy weather: Katrina and the politics of disposability.* Paradigm Publishers, Boulder, Colo., 2006. ISBN 1594513287 9781594513282 1594513295 9781594513299.

[72] Barry Glassner. *The culture of fear: why Americans are afraid of the wrong things.* Basic Books, New York, 2009. ISBN 9780465003365 0465003362.

[73] Good For Her. Feminist porn awards, 2013. URL `http://www.goodforher.com/feminist_porn_awards`.

[74] Jessica Gresko. D.C. jail helps inmates vote, a rarity nationwide, November 2012 (accessed April 24, 2014). URL `http://www.usatoday.com/story/news/politics/2012/11/06/dc-jail-helps-inmates-vote/1685161/`.

[75] John Howard Griffin. *Black like me.* Panther books, London, 1965.

[76] Aimee Groth. This stay-at-home feminist has no interest in 'Having it all', March 2013 (accessed April 5, 2014). URL `http://www.businessinsider.com/new-york-mag-feminist-housewife-story-2013-3`.

[77] Davis Guggenheim, Billy Kimball, Lesley Chilcott, Bill Strickland, Geoffrey Canada, Michelle Rhee, Randi Weingarten, Erich Roland, Bob Richman, Greg Finton, Jay Cassidy, Kim Roberts, Christophe Beck, John Legend, Paramount Vantage, Participant Media, Walden Media, Electric Kinney Films, and Paramount Home Entertainment (Firm). Waiting for "Superman", 2011.

[78] R. E. Guthrie. Reward and punishment. *Psychological Review*, 41(5):450–460, 1934. ISSN 1939-1471(Electronic);0033-295X(Print). doi: 10.1037/h0074245.

[79] Paul Haggis, Hughes Winborne, Mark Isham, Sandra Bullock, Don Cheadle, Matt Dillon, Jennifer Esposito, Brendan Fraser, Terrence Howard, Thandie Ludacris, Newton, Ryan Phillippe, Larenz Tate, Michael Pena, and Lions Gate Home Entertainment. Crash. 2004.

[80] Michael Harrington. *The other America; poverty in the United States.* Macmillan, New York, 1964.

[81] James J Higgins. *An introduction to modern nonparametric statistics.* Brooks/Cole, Pacific Grove, CA, 2003. ISBN 0534387756 9780534387754.

[82] Shere Hite. *The Hite report: women and love : a cultural revolution in progress.* Knopf, New York, 1987. ISBN 0394530527 9780394530529.

[83] bell hooks. *Feminism is for everybody: passionate politics.* South End Press, Cambridge, MA, 2000. ISBN 0896086291 9780896086296 0896086283 9780896086289.

[84] Rob Hopkins. *The transition handbook: from oil dependency to local resilience.* Green, Totnes [England], 2008. ISBN 9781900322188 1900322188.

[85] Houghton Mifflin Company. *The American Heritage dictionary of the English language.* Houghton Mifflin, Boston, 2000. ISBN 0395825172 9780395825174 0618082301 9780618082308.

[86] Housing & Development Board. Public housing in singapore, September 2013 (accessed April 28, 2014). URL http://www.hdb.gov.sg/fi10/fi10320p.nsf/w/aboutuspublichousing?opendocument.

[87] James C. Howell. Preventing and reducing juvenile delinquency, 2009 (accessed April 1, 2014). URL http://www.sagepub.com/books/Book231813?prodId=Book231813&ct_p=toc.

[88] Darrell Huff and Irving Geis. *How to lie with statistics*. Norton, New York, 1954. ISBN 0393052648 9780393052640 039309426X 9780393094268.

[89] HuffPost Education. Dawn hochsprung, sandy hook principal, had implemented new security measures, December 2012. URL http://www.huffingtonpost.com/2012/12/14/dawn-hochsprung-sandy-hook-principal-security-measures_n_2303520.html.

[90] Institute for Sex Research and Alfred C Kinsey. *Sexual behavior in the human female,*. Saunders, Philadelphia, 1953.

[91] John P. A. Ioannidis. Why most published research findings are false. *PLoS Med*, 2(8): e124, August 2005 (accessed March 28, 2014). doi: 10.1371/journal.pmed.0020124. URL http://dx.doi.org/10.1371/journal.pmed.0020124.

[92] Daniel Kalla. *Resistance*. Tor, New York, 2006. ISBN 076535439X 9780765354396.

[93] Daniel Kalla. Prison education: an interdisciplinary examination of prisons, punishment, schools, and education, May 2014.

[94] Zachary Karabell. Ignore the unemployment rate. *Wall Street Journal*, February 2014 (accessed April 27, 2014). ISSN 0099-9660. URL http://online.wsj.com/news/articles/SB10001424052702303496804579364852603962422.

[95] George L. Kelling and James Q. Wilson. Broken windows. *The Atlantic*, March 1982 (accessed April 1, 2014). ISSN 1072-7825. URL http://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/?single_page=true.

[96] Alfred C Kinsey, Wardell Baxter Pomeroy, and Clyde E Martin. *Sexual behavior in the human male*. W.B. Saunders Co., Philadelphia, 1948.

[97] Fay Honey Knopp, Barbara Boward, and Mark O Morris. *Instead of prisons: a handbook for abolitionists*. Prison Research Education Action Project, Syracuse, N.Y., 1976.

[98] Alfie Kohn. *Punished by rewards: the trouble with gold stars, incentive plans, A's, praise, and other bribes*. Houghton Mifflin Co., Boston, 1999. ISBN 0618001816 9780618001811.

[99] Jonathan Kozol. *Amazing grace: the lives of children and the conscience of a nation*. Harper-Perennial, New York, 1996. ISBN 0060976977 9780060976972.

[100] Jane Kroger. *Identity development: adolescence through adulthood*. Sage Publications, Thousand Oaks, Calif., 2007. ISBN 0761929592 9780761929598 0761929606 9780761929604.

[101] Stanley Kubrick, Malcolm McDowell, Patrick Magee, Michael Bates, Warren Clarke, John Clive, Adrienne Corri, Carl Duering, Paul Farrell, Clive Francis, Miriam Karlin, James Marcus, Sheila Raynor, Madge Ryan, John Savident, Philip Stone, Pauline Taylor, Margaret Tyzack, Max L Raab, Si Litvinoff, Wendy Carlos, John Alcott, Bill Butler, John Barry, Anthony Burgess, and Warner Home Video (Firm) Warner Bros. Pictures (1969-). A clockwork orange, 2007.

[102] Thomas S Kuhn. *The structure of scientific revolutions*. University of Chicago Press, Chicago, 1970. ISBN 0226458032 9780226458038 0226458040 9780226458045.

[103] Thomas S Kuhn. *The structure of scientific revolutions*. University of Chicago Press, Chicago, IL, 1996. ISBN 0226458075 9780226458076 0226458083 9780226458083.

[104] LadyAgonyAuntieCurd. Dear modern man versus the feminist "Housewife", May 2012 (accessed April 28, 2014). URL `http://lettersfromladycurd.wordpress.com/2012/05/14/dear-modern-man-versus-the-feminist-housewife/`.

[105] Carl Lerner, James Whitmore, Will Geer, Roscoe Lee Browne, John Howard Griffin, Hilltop Company, and Rhino Home Video (Firm). Black like me, 1992.

[106] Henry M. Levin and Cecilia E. Rouse. The true cost of high school dropouts. *The New York Times*, January 2012 (accessed April 27, 2014). ISSN 0362-4331. URL `http://www.nytimes.com/2012/01/26/opinion/the-true-cost-of-high-school-dropouts.html`.

[107] Steven D Levitt and Stephen J Dubner. *Superfreakonomics: global cooling, patriotic prostitutes, and why suicide bombers should buy life insurance*. William Morrow, New York, 2009. ISBN 9780060889579 0060889578.

[108] Glenn C Loury, Pamela S Karlan, Loic J. D Wacquant, and Tommie Shelby. *Race, incarceration, and American values*. MIT Press, Cambridge, Mass., 2008. ISBN 9780262278577 026227857X 9781435662889 1435662881.

[109] Jay Mathews. Maintaining discipline is key to students' success, but new teachers rarely learn classroom management. *The Washington Post*, page 19, December 2000 (accessed April 20, 2014). URL `http://www.washingtonpost.com/wp-dyn/articles/A23017-2000Dec18.html`.

[110] Melissa McEwan. Shakesville: Trigger by void, March 2009 (accessed March 28, 2014). URL `http://www.shakesville.com/2009/03/trigger-by-void.html`.

[111] Linda Miller, Kren Hess, and Christine Orthmann. *Community Policing: Partnerships for Problem Solving*. Cengage Learning, January 2010. ISBN 1435488687.

[112] Neal E. Miller. Studies of fear as an acquirable drive: I. fear as motivation and fear-reduction as reinforcement in the learning of new responses. *Journal of Experimental Psychology*, 38 (1):89–101, 1948. ISSN 0022-1015(Print). doi: 10.1037/h0058455.

[113] Philip Moeller. Religion makes people HappierBut why? - US news, April 2012 (accessed April 19, 2014). URL `http://money.usnews.com/money/personal-finance/articles/2012/04/12/religion-makes-people-happierbut-why`.

[114] Torin Monahan. *Surveillance in the time of insecurity*. Rutgers University Press, New Brunswick, N.J., 2010. ISBN 9780813549439 0813549434. URL http://site.ebrary.com/id/10393234.

[115] Torin Monahan and Rodolfo D Torres. *Schools under surveillance cultures of control in public education*. Rutgers University Press, New Brunswick, N.J., 2010 (accessed April 6, 2014). ISBN 9780813548265 0813548268. URL http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=311968.

[116] Mark Motivans. Federal justice statistics, 2010. Technical report, U.S. Department of Justice, December 2013 (accessed April 28, 2014). URL http://www.bjs.gov/content/pub/pdf/fjs10.pdf.

[117] Chongmin Na and Denise C Gottfredson. Police officers in schools: Effects on school crime and the processing of offending behaviors. *Justice Quarterly*, 30(4):619–650, 2013. ISSN 0741-8825.

[118] National Institute of Justice and Bureau of Justice Assistance. "Swift and certain" sanctions in probation are highly effective: Evaluation of the HOPE program, February 2012 (accessed April 13, 2014). URL http://www.nij.gov/topics/corrections/community/drug-offenders/Pages/hawaii-hope.aspx.

[119] Alexander Sutherland Neill and Albert Lamb. *Summerhill School: a new view of childhood*. St. Martin's Griffin, New York, 1992. ISBN 0312141378 9780312141370.

[120] Kim E Nielsen. *A disability history of the United States*. 2013. ISBN 0807022047 9780807022047.

[121] Gottfried E Noether and Marilyn Dueker. *Introduction to statistics: the nonparametric way*. Springer-Verlag, New York, 1991. ISBN 0387972846 9780387972848 3540972846 9783540972846.

[122] Gregory Palast. Floridas flawed voter-cleansing program, December 2000 (accessed April 8, 2014). URL http://www.salon.com/2000/12/04/voter_file/.

[123] parentwin. How to be a housewife and a feminist at the same time, July 2013 (accessed April 5, 2014). URL http://www.blogher.com/how-be-housewife-and-feminist-same-time.

[124] Robert Perkinson. *Texas tough: the rise of America's prison empire*. Metropolitan Books, New York, 2010. ISBN 9780805080698 0805080694.

[125] Joan Petersilia. *When prisoners come home: parole and prisoner reentry*. Oxford University Press, Oxford; New York, 2003. ISBN 019516086X 9780195160864.

[126] Becky Pettit. *Invisible men: mass incarceration and the myth of black progress*. Russell Sage Foundation, New York, 2012. ISBN 9780871546678 0871546671 9781610447782 1610447786.

[127] Becky Pettit. Mass incarceration in america: An interview with becky pettit, June 2012 (accessed April 28, 2014). URL http://www.russellsage.org/blog/mass-incarceration-america-interview-becky-pettit.

[128] Elvis Picardo. The unemployment rate: Get real, December 2013 (accessed April 27, 2014). URL http://www.investopedia.com/articles/economics/10/unemployment-rate-get-real.asp.

[129] Kay Pranis, Barry Stuart, and Mark Wedge. *Peacemaking circles: from crime to community*. Living Justice Press, St. Paul, Minn., 2003. ISBN 0972188606 9780972188609.

[130] Prison Gerrymandering Project. Too big to ignore: How counting people in prisons distorted census 2000. Technical report, Prison Policy Initiative, April 2004 (accessed April 19, 2014). URL http://www.prisonersofthecensus.org/toobig/.

[131] Kyle Pyle and Craig Gilmore. Prison town. Comic, The Real Cost of Prisons Project, 2005 (accessed April 27, 2014). URL http://www.realcostofprisons.org/materials/comics/prison_town.pdf.

[132] Tom Rath. *Strengths finder 2.0*. Gallup Press, New York, 2007. ISBN 9781595620156 159562015X.

[133] Diane Ravitch. *The death and life of the great American school system: how testing and choice are undermining education*. Basic Books, New York, 2010. ISBN 9780465014910 0465014917.

[134] Cecil Reynolds, Jane Conoley, Enedina Garcia-Vazquez, Sandra Graham, Peter Sheras, and Russell Skiba. APA zero tolerance task force report, August 2006 (accessed April 5, 2014). URL http://www.apa.org/pubs/info/reports/zero-tolerance.aspx.

[135] Carl R. Rogers. *On becoming a person; a therapist's view of psychotherapy*. Houghton Mifflin, Boston, 1961.

[136] David Rosenbaum. Statisticians charge draft lottery was not random. *New York Times*, January 1970.

[137] Marshall B Rosenberg. *Nonviolent communication: a language of life*. PuddleDancer Press, Encinitas, CA, 2003. ISBN 1892005034 9781892005038.

[138] RT. U.S. prisons hold 10 times more mentally ill people than state hospitals report, April 2014 (accessed April 9, 2014). URL http://rt.com/usa/mental-illness-jail-hospital-281/.

[139] Pasi Sahlberg and Andy Hargreaves. *Finnish lessons: what can the world learn from educational change in Finland?* Teachers College Press, New York, 2011. ISBN 9780807752579 0807752576.

[140] Eugenie Samuel Reich. Dark-matter experiment comes up empty-handed. *Nature*, October 2013 (accessed April 19, 2014). ISSN 1476-4687. doi: 10.1038/nature.2013.14057. URL http://www.nature.com/news/no-sign-of-dark-matter-in-underground-experiment-1.14057.

[141] Allison Samuels. I dont watch Orange is the new black, August 2013 (accessed March 28, 2014). URL http://www.thedailybeast.com/articles/2013/08/19/why-i-don-t-watch-orange-is-the-new-black-or-any-shows-with-black-people-in-prison.html.

[142] Michael Santos. *Inside: life behind bars in America*. St. Martin's Griffin ; Melia, distributor], New York; Godalming, 2007. ISBN 9780312343507 0312343507.

[143] Martin E. P Seligman. *Helplessness: on depression, development and death.* W. H. Freeman, New York, 1992. ISBN 071672328X 9780716723288.

[144] Gene Sharp. *From Dictatorship to Democracy*. Profile, London, 2011 (accessed November 13, 2013). ISBN 9781847657848 1847657842. URL `http://public.eblib.com/EBLPublic/PublicView.do?ptiID=740134`.

[145] Gene Sharp. *How nonviolent struggle works*. 2013. ISBN 9781880813157 1880813157.

[146] Susan M Shaw and Janet Lee. *Women's voices, feminist visions: classic and contemporary readings*. McGraw-Hill Higher Education, Boston, 2009. ISBN 9780073512280 0073512281.

[147] David Shenk. The 32-million word gap, March 2010 (accessed March 28, 2014). URL `http://www.theatlantic.com/technology/archive/2010/03/the-32-million-word-gap/36856/`.

[148] Arun Kumar Singh. *The Comprehensive History of Psychology*. Motilal Banarsidass Publ., 1991. ISBN 9788120808041.

[149] B. F Skinner. *Science and human behavior.* Macmillan, New York, 1953. ISBN 0029290406 9780029290408.

[150] Christina Hoff Sommers. *The war against boys: how misguided feminism is harming our young men*. Simon & Schuster, New York, 2000. ISBN 0684849569 9780684849560 0684849577 9780684849577.

[151] Matthew Sparkes. What is foucault's pendulum? *Telegraph.co.uk*, September 2013 (accessed April 19, 2014). URL `http://www.telegraph.co.uk/technology/google/google-doodle/10317223/What-is-Foucaults-Pendulum.html`.

[152] Claude Steele. *Whistling Vivaldi: how stereotypes affect us and what we can do*. W.W. Norton & Company, New York, 2011. ISBN 9780393339727 0393339726.

[153] Christina Sterbenz. How new york city became safe again, August 2013 (accessed April 1, 2014). URL `http://www.businessinsider.com/criticism-for-giulianis-broken-windows-theory-2013-8`.

[154] Sunzi, Ralph D Sawyer, and Mei-chun Sawyer. *The essential art of war = Sun-tzu ping-fa*. Basic Books, New York, 2005. ISBN 0465072046 9780465072040.

[155] Surveillance Society Network. An introduction to the surveillance society | the surveillance studies network, 2014 (accessed March 19, 2014). URL `http://www.surveillance-studies.net/?page_id=119`.

[156] Emily Swanson. Poll: Few identify as feminists, but most believe in equality of sexes, April 2013 (accessed April 1, 2014). URL `http://www.huffingtonpost.com/2013/04/16/feminism-poll_n_3094917.html#slide=1716233`.

[157] Nassim Nicholas Taleb. *The black swan: the impact of the highly improbable*. Penguin, London, 2008. ISBN 9780141034591 0141034599.

[158] Nassim Nicholas Taleb. *Antifragile: things that gain from disorder*. Random House, New York, 2012. ISBN 9781400067824 1400067820.

[159] James Taylor. As its global warming narrative unravels, the IPCC is in damage control mode, September 2013 (accessed March 29, 2014). URL `http://www.forbes.com/sites/jamestaylor/2013/09/26/as-its-global-warming-narrative-unravels-the-ipcc-is-in-damage-control-mode/`.

[160] Amanda Terkel. Columbine high school had armed guard during massacre in 1999, December 2012 (accessed March 19, 2014). URL `http://www.huffingtonpost.com/2012/12/21/columbine-armed-guards_n_2347096.html`.

[161] Shanti Thakur, Mark Zannis, Judith Gurber-Stitzer, and National Film Board of Canada. Circles, 2007.

[162] Richard H Thaler and Sunstein . *Nudge: improving decisions about health, wealth, and happiness*. Yale University Press, New Haven, Conn., 2008. ISBN 9780300122237 0300122233.

[163] Richard H. Thaler. Opting in vs. opting out. *The New York Times*, September 2009 (accessed March 29, 2014). ISSN 0362-4331. URL `http://www.nytimes.com/2009/09/27/business/economy/27view.html`.

[164] The Sentencing Project. Trends in U.S. corrections. Technical report, The Sentencing Project, Washington, D.C., May 2012.

[165] Beverly B Title. *Teaching peace: a restorative justice framework for strengthening relationships*. Del Hayes Press, Allen, Texas, 2011. ISBN 9780982270646 098227064X.

[166] Jeremy Travis. *But they all come back: facing the challenges of prisoner reentry*. Urban Institute Press, Washington, D.C., 2005. ISBN 0877667500 9780877667506.

[167] Demographic Internet Staff U. S. Census Bureau. A joint effort between the bureau of labor statistics and the census bureau, 2014 (accessed April 27, 2014). URL `https://www.census.gov/cps/`. Current Population Survey (CPS). Main section. Demographic data and reports from the U.S. Census Bureau.

[168] Manza Jeff Uggen, Christopher. Democratic contraction? the political consequences of felon disenfranchisement in the united states. *American sociological review.*, 67(6):777, 2002. ISSN 0003-1224.

[169] Mark S Umbreit, Robert B , Coates, and Boris Kalanj. *Victim meets offender: the impact of restorative justice and mediation*. Criminal Justice Press, Monsey, N.Y., 1994. ISBN 188179802X 9781881798026.

[170] United States and National Commission on Excellence in Education. *A nation at risk: the imperative for educational reform : a report to the Nation and the Secretary of Education, United States Department of Education*. National Commission on Excellence in Education : [Supt. of Docs., U.S. G.P.O. distributor], Washington, D.C., 1983.

[171] U.S. Department of Education. Race to the top fund, March 2014 (accessed April 9, 2014). URL http://www2.ed.gov/programs/racetothetop/index.html. The Race to the Top Fund provides competitive grants to encourage and reward States that are creating the conditions for education innovation and reform.

[172] Deborah J. Vagins. Is race discrimination in school discipline a real problem?, January 2014 (accessed April 13, 2014). URL https://www.aclu.org/blog/racial-justice/race-discrimination-school-discipline-real-problem.

[173] Vox-FRN. QUILTBAG, January 2011. URL http://www.urbandictionary.com/define.php?term=QUILTBAG.

[174] Loic Wacquant. From slavery to mass incarceration. *New Left Review*, (13):41–60, February 2002. ISSN 0028-6060.

[175] Chamont Wang. *Sense and nonsense of statistical inference: controversy, misuse, and subtlety*. Marcel Dekker, New York, 1993. ISBN 0824787986 9780824787981.

[176] Bruce Western. *Punishment and inequality in America*. Russell Sage, New York, 2006. ISBN 0871548941 9780871548948 9780871548955 087154895X.

[177] Richard G Wilkinson and Kate Pickett. *The spirit level: why greater equality makes societies stronger*. Bloomsbury, New York; London, 2011. ISBN 1608193411 9781608193417.

[178] Tim J Wise, Sut Jhally, Jason T Young, David Rabinovitz, and Media Education Foundation. The pathology of privilege racism, white denial & the costs of inequality, 2008.

[179] Jeffrey M Wooldridge. *Econometrics*. Cengage Learning India Pvt. Ltd., New Delhi, 2009. ISBN 9788131509609 8131509605.

[180] Malcolm X and Alex Haley. *The autobiography of Malcolm X*. Ballantine Books, New York, 1992. ISBN 0345379756 9780345379757 0345915038 9780345915030 0345350685 9780345350688.

[181] Dennis Yang. Can we fix the skills gap?, August 2013 (accessed March 31, 2014). URL http://www.forbes.com/sites/groupthink/2013/08/02/can-we-fix-the-skills-gap/.

[182] Karen Zauber. Management tips for new teachers, 2003 (accessed April 20, 2014). URL http://www.nea.org//tools/management-tips-for-new-teachers.html.

[183] Howard Zehr. *Changing lenses: a new focus for crime and justice*. Herald Press, Scottdale, Pa., 1990. ISBN 0836135121 9780836135121.

[184] Howard Zehr. *The little book of restorative justice*. Good Books, Intercourse, PA, 2002. ISBN 1561483761 9781561483761.

[185] Howard Zehr and Ali Gohar. *The Little Book of Restorative Justice*. Unicef, http://www.unicef.org/tdad/littlebookrjpakaf.pdf, October 2003 (accessed April 28, 2014).