

Intersections between Issues of  
Diversity, Equity, and Inclusion and  
the Field of Statistics:  
Reading Packet for the Fall 2022  
Elementary Statistics I Sections  
Taught by Dr. Jana Asher





# Fall 2022 STAT 152 (HIP-C, DEI)

## Graphical Syllabus

### 1. Data Sources

- Observations
- Experiments
- Censuses
- Surveys
- Simulations
- Administrative /Found Data

### 2. Data Visualization

- Bar Graphs
- Pie Charts
- Histograms
- Scatterplots
- Rose plots
- Time series

### 3. Descriptive Statistics

- Central Tendency (mean, median, mode)
- Spread (standard deviation, range)
- Position (quantiles)
- Relationship (correlation)

### 4. Probability

- Addition Rule
- Multiplication Rule
- Mutually Exclusive
- Conditional Probability
- Bayes Theorem
- Probability Functions
- Binomial Distribution
- Normal Distribution
- Law of Large Numbers
- Distribution of Sample Means
- Central Limit Theorem

### 5. Inference

- Confidence Intervals
- Hypothesis Testing
- Linear Regression

### 1. Data Sources and Diversity, Equity, and Inclusion (DEI):

**Implicit Bias –**  
The Mismeasure of Man  
**Structural Oppression –**  
Tuskegee Syphilis Study  
Japanese Internment  
Differential Undercount

### 2. Data Visualization and DEI: Diversity in the Field

**Structural Oppression–**  
Hold your ground/Florida

### 3. Descriptive Statistics and DEI:

**Structural Oppression–**  
Bias and Assumptions  
Association/Causation

### 4. Probability and DEI:

**Intersectionality–**  
Statistical Independence

### 5. Statistical Inference and DEI:

**Structural Oppression–**  
Gender in Data  
Eugenics  
Public Policy

**STATISTICS ARE THE  
HEART OF DEMOCRACY**

SIMEON STRUNSKY

### Course Grade

Homework: 20%  
In-Class Laboratories: 20%  
Participation in Discussions: 5%  
Reflections: 15%  
Tests: 20%  
Final Project: 20%

### Assessments/Deadlines

#### Week 1:

In-class Lab 1 on 1/21

#### Week 2:

Finish Reading 1 by 1/24  
Reflection 1 due 1/26  
Homework 1 due 1/28  
In-class Lab 2 on 1/28

#### Week 3:

Finish Reading 2 by 1/31  
Reflection 2 due 2/2  
HW 2 due 2/4  
In-class Lab 3 on 2/4

#### Week 4:

Finish Reading 3 by 2/7  
Reflection 3 due 2/9  
HW 3 due 2/11  
In-class Lab 4 on 2/11

#### Week 5:

**TEST 1**, covering Data Sources, Data Visualization, and Descriptive Statistics on 2/16

#### Week 6:

Finish Reading 4 by 2/21  
Reflection 4 due 2/23  
HW 4 due 2/25  
In-class Lab 5 on 2/25

#### Week 7:

Finish Reading 5 by 2/28  
Reflection 5 due 3/3  
HW 5 due 3/4  
In-class Lab 6 on 3/4

#### Week 8:

Finish Reading 6 by 3/7  
Reflection 6 due 3/9  
HW 6 due 3/11  
In-class Lab 7 on 3/11

#### Week 9: SPRING BREAK!

#### Week 10:

**TEST 2**, covering Probability on 3/23

#### Week 11:

Finish Reading 7 by 3/28  
Reflection 7 due 3/30  
HW 7 due 4/1  
In-class Lab 8 on 4/1

#### Week 12:

Finish Reading 8 by 4/4  
Reflection 8 due 4/6  
HW 8 due 4/8

#### Week 13:

Finish Reading 9 by 4/11  
Reflection 9 due 4/13  
HW 9 due 4/15  
In-class Lab 9 on 4/15

#### Week 14:

Finish Reading 10 by 4/18  
Reflection 10 due 4/20  
HW 10 due 4/22  
In-class Lab 10 on 4/22

#### Week 15:

Finish Reading 11 by 4/25  
Reflection 11 due 4/27  
HW 11 due 4/29

#### Week 16:

**TEST 3**, covering Statistical Inference on 5/2  
**Final Project Presentations** on 5/4  
**Final Paper** due on 5/6

### List of Readings and Student Learning Outcomes Overall and by Unit

DEI Designation Course Learning Outcome	Major Outcome (Mathematics)	University Outcome
CO 10. Students will articulate the connection between the statistical concepts of demographics, independence, assumptions, bias, causation, correlation, data visualization, and hypothesis testing and the DEI concepts of social identity, diversity, intersectionality, marginalization, discrimination, implicit bias, structural privilege, structural oppression, cultural competence, and social justice.	5. Demonstrate professional integrity and accountability in the use and communication of mathematics.	5. Develop as a whole person: (a) Acquire and apply knowledge and skills in the major and profession, including soft skills. (b) Make connections beyond traditional disciplinary boundaries. 9. Develop a worldview that acknowledges diversity and global interdependence: (a) Understand the importance of diverse experiences, cultures, and identities. (b) Understand the ways that group and individual inequalities and interactions impact self and society. (c) Apply multiple perspectives to address local, regional, global, and cultural issues.
CO 11. Students will cite examples of how an individual's social and cultural identity can lead to biased scientific conclusions that exacerbate structural oppression.		8. Act ethically: (a) Respect the range of ethical perspectives. (b) Understand their own values and principles. (c) Recognize the consequences and impacts of their actions on others.
CO 12. Students will analyze examples of poorly-designed/unethical and well-designed/ethical methods for data collection and analysis in the context of structural oppression and/or privilege.		9. Develop a worldview that acknowledges diversity and global interdependence: (a) Understand the importance of diverse experiences, cultures, and identities. (b) Understand the ways that group and individual inequalities and interactions impact self and society. (c) Apply multiple perspectives to address local, regional, global, and cultural issues.
CO 13. Students will articulate the connections between demographic data collection and human diversity and how statistics can create or exacerbate structural oppression of demographic groups.		

#### Unit 1: Concepts of Diversity, Equity, and Inclusion in the Field of Statistics

**UO11:** Through readings, group discussion and reflection, students will state definitions for DEI concepts including diversity, equity, inclusion, structural oppression, implicit bias, microaggressions, privilege, power, and marginalization. (CO 10)

**UO12:** Through readings, group discussion and reflection, students will determine ways in which the field of statistics contributes to inequities or equity in society. (CO 10)

#### Unit 1 Readings

- Structural Oppression: 5 Concepts, 1 Theory
- How to Help Advocate for Justice, Equity, Diversity, and Inclusion
- Addressing Implicit Bias Among Women Statisticians and Data Scientists
- A conversation about implicit bias
- Power and Privilege: Reshaping the Opportunity Structure for Equitable Leadership in Statistics and Data Science
- At the Rise of JEDI: Lessons Learned from Fall of the Jedi Order in Star Wars

**Unit 1 Discussion Prompts:**

1. Several concepts related to diversity, equity, and inclusion are introduced through these readings, including structural oppression, implicit bias, microaggressions, privilege, power, and marginalization. From the readings, what do these terms mean?
2. In these readings, the relationship between the field of statistics and DEI concepts are presented mostly through the lens of female statisticians that have experienced inequities in their careers. Can you think of other ways that DEI concepts are relevant to the field of statistics?

**Unit 1 Reflection Prompts:**

1. Think about the DEI concepts discussed this week and how they relate to your own life. Can you think of a time when you felt privileged or marginalized? Do you feel like you are in a position of privilege or of marginalization in your own life? Why?
2. Have you ever had someone obtain data from or about you that you didn't want to share? Where did that happen, and for what purpose? What type of power did that information give to the person that had it?

**Unit 2: Study Design**

**UO21:** Through readings, group discussion and reflection, students will explain through specific real-life examples how individual racism and discrimination can lead scientists to make unethical decisions around data that contribute to structural oppression of marginalized groups. (CO 11)

**UO22:** Through readings, group discussion and reflection, students will analyze how power that scientists hold in society interacts with societal stereotypes about marginalized groups. (CO 11)

**Unit 2 Readings**

- a. Racism and Research: The Case of the Tuskegee Syphilis Study
- b. A New Look at an Old Problem: Improving Diversity, Equity, and Inclusion in Scientific Research
- c. The Mismeasure of Science: Stephen Jay Gould versus Samuel George Morton on Skulls and Bias
- d. Mismeasuring Gould
- e. The History of the U.S. Census and the Undercount
- f. The Census and the Japanese "Internment": Apology and Policy in Statistical Practice

**Unit 2 Discussion Prompts:**

1. What are some examples of structural violence against African-Americans and Japanese-Americans from these readings? How are these examples related to the field of statistics?
2. The fifth reading, from the book *Who Counts?* The politics of census-taking in contemporary American, introduces the idea of the differential undercount. What is a differential undercount, and how is it related to the DEI concept of structural oppression?

**Unit 2 Reflection Prompts:**

1. In this reflection we will be exploring our social identities as we consider how society is structured. Complete the following chart about your own identity. You do not have to share this identity in your written reflection.

<b>Domain</b>	<b>My Identity</b>	<b>At Slippery Rock University, this identity gives me a position of: Privilege (P), Marginalization (M)</b>	<b>In the place I grew up, this identity gave me a position of: Privilege (P), Marginalization (M)</b>
<b>Race</b> (e.g., white, black, biracial)			
<b>Ethnicity</b> (e.g., German, Polish, Cuban, Hispanic, Japanese)			
<b>Gender identity/ expression</b> (e.g., woman, man, nonbinary, transgender male, transgender female)			
<b>Sexual orientation</b> (e.g., homosexual, heterosexual, bisexual, pansexual)			
<b>Religion</b> (e.g., Catholic, Protestant, Jewish, Muslim)			
<b>Dis/ability</b> (e.g., able-bodied, autistic, ADHD, deaf)			
<b>First Language</b> (e.g., English, Spanish)			
<b>First Generation Status</b> (first generation in college in nuclear family, not first generation in college)			
<b>Socio-Economic Status</b> (as you feel it is best defined – use of governmental assistance programs such as WIC, Medicaid, or SNAP; middle class, working class, etc.)			
<b>Other:</b> (you define)			

After mapping your social identity, think about the readings and the experiences of the people that were or are discriminated against. How are their social identities similar or different than yours?

2. How did the readings make you feel about the structural inequities in society that result from flawed data collection processes?

3. What responsibility do you feel toward making sure scientific and governmental data collection efforts are equitable?

### **Unit 3: Data Visualization**

**UO31:** Through readings, group discussion and reflection, students will analyze how a statistician's social identity translates into biased understanding of the history of statistical thought. (CO 11)

**UO32:** Through readings, group discussion and reflection, students will critique the dominant narrative about the history of data visualization by comparing and contrasting euro-centric analysis of the history of statistical thought and more diverse understandings of the history of statistical thought. (CO 11)

#### **Unit 3 Readings**

- a. A Brief History of Data Visualization
- b. REVIEW: Data + Diversity-Exploring the Data Visualizations of W.E.B. Du Bois
- c. Beyond Nightingale: Being a Woman in Data Visualization
- d. Asia's data scene deserves greater attention. That's why we are starting a movement.

#### **Unit 3 Discussion Prompts:**

1. In the first reading, the history of data visualization is given. How much of the history of data visualization is centered in Western Europe and the United States?
2. The last three readings explore experiences of scientists that are not white and male. How does each of these scientists experience tokenism or structural violence in their career? How are they resisting the cultural norm?

#### **Unit 3 Reflection Prompts:**

1. What are some of the ways that a Euro-centric focus on the history of data visualization affects and distorts our conceptualization of data visualization today?
2. What are some of the ways that modern statisticians are working to rewrite the history of data visualization and make it more inclusive?

### **Unit 4: Graphical Distortion**

**UO41:** Through readings, group discussion and reflection, students will evaluate misleading graphical summaries of data with an equity lens to explore how graphical summaries can exacerbate structural oppression. (CO 12)

**UO42:** Through readings, group discussion and reflection, students will identify best practices for promoting equity through appropriate data visualization techniques. (CO 12)

#### **Unit 4 Readings**

- a. Be Mindful of Misrepresenting Data in Visualization
- b. Do No Harm Guide: Applying Equity Awareness in Data Visualization

#### **Unit 4 Discussion Prompts:**

1. In the examples given in the first reading, how do the graphical mistakes made reflect the implicit biases of the creators? How does the author equate the creation of misleading graphics to power and privilege?
2. What techniques do the researchers at The Urban Institute recommend to create equity within and through data visualization? Do you agree with their assessments?

#### **Unit 4 Reflection Prompts:**

1. Thinking in terms of your major and your future career, which types of data misrepresentations in statistical graphics do you think you will be most likely to encounter?
2. What steps can you take in your future career to ensure that any statistical graphics you create contribute to a more equitable society?

### **Unit 5: Descriptive Statistics**

**UO51:** Through readings, group discussion and reflection, students will analyze how unintentional and intentional misinterpretation of descriptive statistics can influence public debate about DEI-related issues. (CO12)

#### **Unit 5 Readings**

- a. Faulty Interpretation
- b. 7 Statistics That Will Change How You View Racism (9 pages)
- c. No, the United States is not Systemically Racist (2 pages)

#### **Unit 5 Discussion Prompts:**

1. What are some of the misuses of statistics listed in the first reading? Are they all conscious misuses? Could implicit bias make these misuses more likely?
2. In the last two readings, the authors use descriptive statistics to argue opposing viewpoints as to whether the U.S. is racist or not. Can you find any of the misuses listed in the first article? Which author makes the best argument and why?

#### **Unit 5 Reflection Prompts:**

1. You have had a chance to read and discuss differing viewpoints on whether the U.S., as a country, is racist. What do you now think, and why?
- 2a. If you think that the U.S. is racist, what do you think your personal responsibility is as a citizen of the country?
- 2b. If you think that the U.S. is not racist, how do you explain the existing socioeconomic and educational divide between different racial groups?

### **Unit 6: Correlation and Causation**

**UO61:** Through readings, group discussion and reflection, students will critique causal arguments related to race in the United States by analyzing the relationship between the author's social identity and the appropriateness of the use of statistical concepts within the argument. (CO 11)

**U062:** Through readings, group discussion and reflection, students will connect cross-classification and Simpson's Paradox to concepts of intersectionality. (CO 13)

### Unit 6 Readings

- a. Policing, Segregation, and Causation vs. Correlation
- b. Race in America: A Conversation
- c. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist policies.
- d. Simpson's Paradox Explained, or when Facts aren't really Facts.

### Unit 6 Discussion Prompts:

1. In the first reading, the correlation between racial segregation and police violence is interpreted as causal in both directions by different scholars: Schecter believes segregation leads to police violence, and Axel-Lute believes that police violence is necessary for racial segregation to be maintained. Can either of these points of view be proved with the statistics collected? Does Axel-Lute make a convincing argument? Why/why not?
2. In the second reading, Wilkes challenges the "popular" narrative about race in the U.S. and asserts that causality and correlation are being confused when exploring negative experiences of African Americans. What is Leah's response? How does her response demonstrate the power of combining qualitative information with quantitative data to make causal arguments?
3. What is intersectionality? What is Simpson's paradox? How do you think Simpson's paradox is related to intersectionality?

### Unit 6 Reflection Prompts:

1. What arguments do the authors of the readings make about the reasons for inequities between races in United States society? Do you agree with one of the authors? Why/why not?
2. Simpson's paradox shows us that we can see evidence of discrimination at one level of data disaggregation and have that evidence reverse at another level of data disaggregation. For example, we might note that there is no evidence of discrimination against a particular racial group across an entire college, but there is evidence within certain college majors, or vice versa. How does this fact reflect the concept of intersectionality?

## Unit 7: Probability

**U071:** Through readings, group discussion and reflection, students will identify, through describing real examples, how statistical knowledge can be a source of power. (CO 13)

**U072:** Through readings, group discussion and reflection, students will examine how current statistical methods are not sufficient to describe the complexities of intersectionality. (CO 10)

### Unit 7 Readings

- a. Power
- b. Statistics for Equity: Capturing, Not Masking, Intersectional Dynamics in Data
- c. Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods

**Unit 7 Discussion Prompts:**

1. In the first reading, Su explains the power of mathematics through a probability example. What type of power does Su believe mathematicians yield? How does she contrast that power with coercive power? Which type of power is associated with equity? With structural oppression?
2. What is the main argument about the ability of statisticians to use data to model intersectionality? How is this argument related to the addition rule and the multiplication rule we learned in class?
3. What issues are current researchers encountering when attempting to understand intersectionality using quantitative methods?

**Unit 7 Reflection Prompts:**

1. Think about your own experiences both in the classroom and outside of it with statistics. What causes statisticians, or people that understand statistics well, to yield greater power than those that don't understand those methods?
2. Think about your social identity as you defined it in the intersectionality map you created during Unit 2 of these reading assignments. Do you think your identities have an additive effect, a multiplicative effect, or a more complex effect on your status of privilege or marginalization in society? Why?

**Unit 8: Gender and Data**

**U081:** Through readings, group discussion and reflection, students will define different sexual orientations and gender identities and determine which orientations and identities are better represented in current data collection and analysis practice in the United States. (CO 13)

**U082:** Through readings, group discussion and reflection, students will connect how data collection and analysis practice on gender expression and sexual orientation can lead to structural privilege or oppression in United States society. (CO 13)

**Unit 8 Readings**

- a. Collect Data on Sexual Orientation and Gender Identity in Health Surveys
- b. New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys
- c. Collection of Data on Sex, Sexual Orientation, and Gender Identity by U.S. Public Health Data and Monitoring Systems, 2015–2018

**Unit 8 Discussion Prompts:**

1. In the first two readings, the authors explain barriers related to understanding marginalization of the LGBTQIA+ population. What are those barriers? Why would removing those barriers make equity for the LGBTQIA+ population more likely?
2. Why could problems in data collection, as described in all three readings, be considered a form of structural oppression?

**Unit 8 Reflection Prompts:**

1. Many individuals in the United States are uncomfortable thinking about or supporting individuals of different sexual orientations and gender identities. What do you think is the responsibility of individual people to support the rights of non-cis-gender and heteronormative individuals in United States society?



2. Think about your own beliefs about gender expression and sexual orientation. Are they different than your parents' beliefs? Have your beliefs changed since you became a student at Slippery Rock University? If so, how and why have they changed?

### **Unit 9: Statistics and Public Policy**

**U091:** Through readings, group discussion and reflection, students will explain how the incorrect use or collection of statistics by governments can lead to structural oppression of marginalized groups. (CO 10, CO 12)

**U092:** Through readings, group discussion and reflection, students will explain how the incorrect application of hypothesis testing can exacerbate structural inequalities in society. (CO 10, CO 12)

#### **Unit 9 Readings**

- a. The Body Politic: Governments and Politicians
- b. How Misused Statistics can Harm Democracy (2 pages)
- c. Statistics and the Millennium Development Goals

#### **Unit 9 Discussion Prompts:**

1. In what ways are individual liberties curtailed and inequities preserved when governments refuse to collect or misinterpret statistics?
2. The last reading outlines the process of Lot Quality Assurance Sampling used by Valadez. How is Valadez's method flawed, and what does that mean in terms of global health inequities? What do the authors of the article suggest should be done instead?

#### **Unit 9 Reflection Prompts:**

1. Think of and write down an example of how a hypothesis test, if set up with the wrong null hypothesis, could exacerbate structural inequalities in society. Try to have your example not be about vaccinations or similar medical processes, or quality control of meat or similar processes. What is another area of society where an incorrectly set up hypothesis test could cause significant structural oppression?

### **Unit 10: Statistics and Eugenics**

**U101:** Through readings, group discussion and reflection, students will analyze how the social identities of the originators of modern statistical inference led them to also be proponents of eugenics. (CO 11)

**U102:** Through readings, group discussion and reflection, students will describe how incorrect scientific assumptions led to the inappropriate application of statistical methods to make eugenics arguments. (CO 11)

#### **Unit 10 Readings**

- a. Is Statistics Racist?
- b. Statistics, Eugenics, and Me
- c. The outstanding scientist, R.A. Fisher: his views on eugenics and race

**Unit 10 Discussion Prompts:**

1. In 2000, the American Statistical Association retired the Fisher Lecture, the most prestigious speaking engagement in the statistics community, due to growing concerns about Fisher's connections to eugenics. Was that the correct decision? Why or why not?
2. The authors of the last reading argues that Fisher is not racist. Do you agree? What mistakes did they make in their arguments?

**Unit 10 Reflection Prompts:**

1. Do you think that the ASA should have retired the Fisher Lecture? Why or Why Not?
2. What incorrect assumptions did the founders of modern statistical methods make about the relationship between intelligence and social standing, and the relationship between intelligence and race?

**Unit 11: Data for Social Justice: Meet Your Professor**

**U111:** Through readings, group discussion and reflection, students will explain how statistics as a field can be a tool for social justice. (CO 12, CO 13)

**U112:** Through readings, group discussion and reflection, students will list specific approaches to data collection and analysis that promote equity. (CO 12, CO 13)

**Unit 11 Readings**

- a. Evidence-Based Documentation of Gender-Based Violence

**Unit 11 Discussion Prompts:**

1. How does this reading demonstrate the relationship between statistics and social justice?
2. What methods do the authors advocate to reduce trauma, share power, and increase equity during the data collection process?

**Unit 11 Reflection Prompts:**

1. Think over everything you have learned about the connection between statistics—the collection, organization, analysis, and dissemination of data—and its connections to building a more equitable society. What have you learned about your own role in this process? How can you, as a professional and as a private citizen, help make the world more equitable through your interactions with data?
2. Over this semester, what did you learn about statistics as a field and its connection to civil society?

# Reading Unit 1: Concepts of Diversity, Equity, and Inclusion in the Field of Statistics

(Estimated Reading Time: 1 hour,  
Complete Reading by 8am 01/24/2022)

**UO11:** Through readings, group discussion and reflection, students will state definitions for DEI concepts including diversity, equity, inclusion, structural oppression, implicit bias, microaggressions, privilege, power, and marginalization. (CO 10)

**UO12:** Through readings, group discussion and reflection, students will determine ways in which the field of statistics contributes to inequities or equity in society. (CO 10)

1. Blair, D. (May 6 2016) Structural Oppression: 5 Concepts, 1 Theory. [anemoiadotxyz.wordpress.com/2016/05/06/structural-oppression-5-concepts-1-theory/](http://anemoiadotxyz.wordpress.com/2016/05/06/structural-oppression-5-concepts-1-theory/) (accessed 2022-01-08).
2. Sellers, K. F. (2021). How to Help Advocate for Justice, Equity, Diversity, and Inclusion. *Amstat News*, 532, 9.
3. Sellers, K. F., Benn, E. K., Garcia, M., & Kellam, M. (2017). Addressing Implicit Bias Among Women Statisticians and Data Scientists. *CHANCE*, 30(2), 38-41.
4. Golbeck, A. L., Ash, A., Gray, M., Gumpertz, M., Jewell, N. P., Kettenring, J. R., ... & Gel, Y. R. (2016). A conversation about implicit bias. *Statistical Journal of the IAOS*, 32(4), 739-755.
5. Benn, E. (2021). Power and Privilege: Reshaping the Opportunity Structure for Equitable Leadership in Statistics and Data Science. *Leadership in Statistics and Data Science: Planning for Inclusive Excellence*, 19.
6. Liao, S-M. (2022). At the Rise of JEDI: Lessons Learned from Fall of the Jedi Order in Star Wars. *Amstat News*, 535, 12-13

# Reading Unit 1:

## Discussion Questions

1. Several concepts related to diversity, equity, and inclusion are introduced through these readings, including structural oppression, implicit bias, microaggressions, privilege, power, and marginalization. From the readings, what do these terms mean?
2. In these readings, the relationship between the field of statistics and DEI concepts are presented mostly through the lens of female statisticians that have experienced inequities in their careers. Can you think of other ways that DEI concepts are relevant to the field of statistics?

# Reading Unit 1:

## Reflection Questions

1. Think about the DEI concepts discussed this week and how they relate to your own life. Can you think of a time when you felt privileged or marginalized? Do you feel like you are in a position of privilege or of marginalization in your own life? Why?
2. Have you ever had someone obtain data from or about you that you didn't want to share? Where did that happen, and for what purpose? What type of power did that information give to the person that had it?

# Structural Oppression: 5 Concepts, 1 Theory<sup>1</sup>

Posted on May 6, 2016 by Danielle Blair

*Someone who does not see a pane of glass does not know that he does not see it. Someone who, being placed differently, does see it, does not know the other does not see it. When our will finds expression outside ourselves in actions performed by others, we do not waste our time and our power of attention in examining whether they have consented to this. This is true for all of us. Our attention, given entirely to the success of the undertaking, is not claimed by them as long as they are docile... Rape is a terrible caricature of love from which consent is absent. After rape, oppression is the second horror of human existence. It is a terrible caricature of obedience. —Simon Weil, retrieved from “The Five Faces of Oppression,” p. 39.*

It can be difficult to grapple with the idea of oppression as a contemporary social construct, especially in a society where oppression, as a word, is taboo—rejected by social customs and expectations. As Sandra Hinson and Alexa Bradley (1) explain in their article, “A Structural Analysis of Oppression,” many people associate the word “oppression” with distant times and places:

*It is what brutal dictators and totalitarian governments do to their subjects or to the people they have conquered. People do not think of oppression as something that happens in open and democratic societies, partly because they associate oppression with an ‘intent’ to oppress.*

In spite of this cultural refusal, oppressive conditions continue to exist in “liberal, democratic societies” (Hinson and Bradley 1). The unintentional nature of contemporary oppression does not combat its existence as a cloth “woven into the fabric of our major economic, political and cultural institutions” (Hinson and Bradley 1); instead, it allows oppression to thrive in its five forms: exploitation, marginalization, powerlessness, cultural dominance and violence (Young 48).

## Who can be oppressed?

According to Iris Marion Young, author of the article, “The Five Faces of Oppression,” it is social groups that are subjected to systemic oppression (41). Further according to Young, “A social group is a collective of persons differentiated from at least one other group by cultural forms, practices, or way of life” (41). Members of a social group identify with one another because of their similarities, in experience or in culture, “which prompt them to associate with one another more than with those not identified within the group, or in a different way” (Young 41). As an “expression of social relations,” a group may only exist in relation to one other group, or more than one other group, and it is through these group-group interactions that individual group identity is formed and strengthened (Young 41). This sense of identity, rather than the attributes that inspired the collectivism of the group, defines the social group within the community (Young 44).

With this in mind, the inter-relatedness of social groups does not require that an oppressed group have a correlating oppressive group; in fact, as young notes in her article, “While structural

---

<sup>1</sup> From <https://anemoiadotxyz.wordpress.com/2016/05/06/structural-oppression-5-concepts-1-theory/>

oppression involves relations among groups, these relations do not always fit the paradigm of conscious and intentional oppression of one group by another” (41).

### What are the five types of structural oppression?

The five types of structural oppression, as they are listed by Hinson and Bradley (2), are:

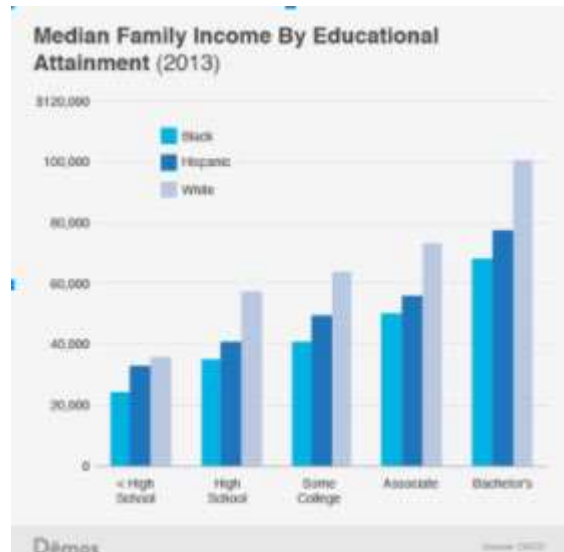
1. Exploitation
2. Marginalization
3. Powerlessness
4. Cultural Dominance
5. Violence

#### *Exploitation*

*This world is ill divided. Them that work the hardest are the least provided.* —English Folk Song, Early 19th Century, retrieved from “A Structural Analysis of Oppression,” p. 2.

In a market economy such as that of the United States, labor is the most profitable commodity for employers (Hinson and Bradley 2). Employers want to keep the surplus “that results from the difference between workers’ productivity and their wages,” which is directly opposed to the goal of the worker: to receive increased pay for labor provided (Hinson and Bradley 2). This means that not only do employers and employees have drastically different perceptions of fair wages, but there is “some degree of exploitation” built inevitably into the employer-employee relationship (Hinson and Bradley 2). According to Hinson and Bradley, “In any society, the extent of the gap between the wealthy owners and the masses of working people is an indication of the degree of exploitation that exists in that society” (2). Also according to Hinson and Bradley, this oppression is particular to race for the following reasons:

- Historically, race-specific exploitation has existed within the capitalist system in the U.S, and elsewhere
- Capitalism seeks to keep a segment of the labor market stuck in, or desperate for, low-paying, low-skill jobs. For historic reasons, people of color make up the bulk of this segment of the labor market.
- Race-based segmentation of workers continues to make it harder for workers of color to get higher paying, higher skilled jobs.
- For many immigrant communities, social isolation and invisibility reinforces race-specific exploitation.
- Discrimination in other spheres, such as housing and education, ensures the continuation of race-based labor market segmentation



- As better-paying jobs become more scarce, and as competition intensifies, people of color with good jobs experience more resentment from white workers who think they got the job through 'affirmative action. (2)

### *Marginalization*

Labor is a commodity, meaning that individuals who perform labor for employers must first market themselves to those employers. This is more difficult for some than others; as Hinson and Bradley describe, "Some segments of the population do not possess skills, attributes or characteristics that employers are seeking. For the most part, they are shut out of the labor market" (3). These individuals, however, should not be regarded only as individuals; they tend to be part of prominent, minority social groups. These groups may include the involuntarily unemployed, the elderly, individuals with physical or intellectual disabilities and individuals who have not received a proper education academically nor occupationally (Hinson and Bradley 2). According to Hinson and Bradley, it is these aforementioned groups that experience marginalization—stigmatization, labelling as "non-productive members of society" despite their responsibilities and successes (2).



In terms of race, youth of color who are unable to obtain a first job are marginalized in ways that affect their future employment opportunities and quality of life (Hinson and Bradley 2). Native Americans may be marginalized due to their high rates of unemployment and "limited opportunities to develop marketable skills" (Hinson and Bradley 2). People who have been convicted of crimes, whom are disproportionately people of color, are likely to be marginalized (Hinson and Bradley 2). Due to workers displacement and the lessening availability of high-paying jobs, workers of color face marginalization from coworkers who assume that they have been hired to fill a diversity quota, and colored workers who can not find jobs in general are denied access to the development opportunities that would ease marginalization (Hinson and Bradley 2).

### *Powerlessness*

Powerlessness is the first aspect of oppression to involve status—societal perceptions of respectability on top of class distinction and autonomy in the workplace. According to Hinson and Bradley, the majority of workers experience powerlessness when they are consistently uninvolved in decisions that "affect the conditions of their employment, and, beyond that, conditions of their lives" (3). According to Hinson and Bradley (3), professional workers, as opposed to employees of the working class, are more likely to have a higher status, "more relative social as well as economic power," due to the following:

- Knowledge, expertise and opportunities to use these on the job and in their daily lives, as well as opportunities to expand them.
- Autonomy, which means they have a voice in the conditions of their employment. They supervise others, and many opportunities to exercise their own judgment and to make significant decisions.

- Social respectability, which means that, on the job and in life in general, professionals enjoy a high social status. Their opinions are sought after and listened to. They are seen to be in control of their lives.

The concept of powerlessness is more often applicable to individuals of color than to white individuals, making race a pertinent factor in this oppression. The majority of people of color are required to prove their respectability, socially and politically; whereas, for white individuals, respectability is assumed or granted (Hinson and Bradley 3). This generally and automatically lessened social respectability “translates into having less political power” in a democracy where professionals, in lieu of blue-collar workers, have the least restricted access to political institutions and politicians (Hinson and Bradley 3). With this in mind, it is uncommon for professionals to recognize their privilege regarding political access, until or unless a professional is to lose his or her status (Hinson and Bradley 3).

### *Cultural Dominance, Imperialism*

Cultural dominance is the first of the terms discussed here to surpass individuals’ class and status distinctions and to encompass the social and political histories associated with the groups in which these individuals are involved. Cultural dominance in particular, as described by Hinson and Bradley (4),

*[Cultural dominance or imperialism] refers to the way that one group’s experiences, cultural expressions and history are defined as superior to all other groups’ experiences and histories. It is not necessary for anyone to say: “my group’s culture is superior;” it simply has to be treated as universal — representing the best in all of humanity. It is considered ‘normal,’ which means that all others are either ‘strange,’ or ‘invisible’ or both.*

In the case of cultural dominance, the dominant culture is continually reinforced by its socially, culturally prioritized position within the community. According to Hinson and Bradley, “members of the culturally dominant group tend to control the means of interpreting, producing and reproducing cultural goods and products: art, music, literature, film, etc.” (4). This indicates that not only does the culturally dominant group already maintain a heightened position, but that those whom are culturally different are outsiders, exotic, the Other—inferior. Hinson and Bradley note: “And the cultural differences that the dominant group sees in others are easily ascribed to physical variation, such as skin color, ethnicity, accents, gender, sexual identities, etc.” (4).

This leads to a plethora of difficulties for individuals, or social groups, who are considered outside the realm of normality (outside the culturally dominant group). Members of minority groups often “feel ‘marked out’ as different,” seen as representing the entire group in lieu of being judged as individuals (Hinson and Bradley 4). This can be observed in several aspects of American culture. For example, the American media tends to vilify Muslims, forcing upon them an assumed terrorist-identity based on the actions of individual Muslims throughout recent years; wherein, individuals who are white and Christian who have committed mass murders, involving suicide bombings, firearms and bio-terrorism tactics, are infrequently called out as terrorists, let alone called out as terrorists and used as a judgement-point for the entire white, Christian population.





In other situations, individuals who are part of a less dominant group may “feel invisible” because their “expressions and experiences are not represented” (Hinson and Bradley 4). Further according to Hinson and Bradley (4):

*All of this gets internalized: you [the minority] look at yourself through the eyes of the dominant group. You struggle against stereotypes and the limits that are placed on you. At the same time, you may feel a deeper connection with members of your cultural and social group, and you want to lift up the rich and meaningful expressions that you and members of your group create and experience. Those who fit more neatly within the mainstream culture also miss out — they lose opportunities to know more about, connect with, people who are different from them. They lose some of the richness of the human experience.*

The greatest injustice of cultural dominance, as is noted by Hinson and Bradley, may be the tendency of the dominant group to “impose its own interpretations of social life upon all others” (4). What is not dominant is different, exotic and less influential; therefore, cultural and economic investments often rely heavily on the influence of the dominant culture (Hinson and Bradley 4). This means that “how we value some neighborhoods, cities and regions over others, whether we see certain uses of public funds as ‘good investments’ or ‘bad investments,’ and whether we value public education enough to invest in all children or just some children” relies on the receiving population’s relation to the cultural imperial, reinforcing marginalization and powerlessness (Hinson and Bradley 4).

### *Violence*

Violence is the fifth and final of the universal concepts of structural oppression. Across the globe, but also within the United States of America, specifically, there are people who “live with fear of random attacks that are meant to humiliate and/or destroy them” because of the social groups with which they identify (Hinson and Bradley 4). The history of the United States, in particular, is rich with instances in which violence has been used to undermine or oppress a specific social group, including violent, state-sanctioned racial segregation (Hinson and Bradley 4). In their article, Hinson and Bradley (4) offer further examples of violence used as a form of humiliation:



- Police brutality against Black and Latino men;
- The way in which rape and sexual harassment keep women vulnerable;
- Attacks on Muslims, or people assumed to be Muslim, especially since 9/11;
- Hate crimes against gays, lesbians and transgendered people;
- Attacks on immigrants at day-labor gathering places, and the constant threat of workplace raids.

Outright violence, however, is not the only threat to Other individuals; ongoing harassment—verbal, sexual, racial profiling—is equally effective in creating a submissive Other and often “carries with it the threat of physical attack” (Hinson and Bradley 4). Violence does not have to be fighting

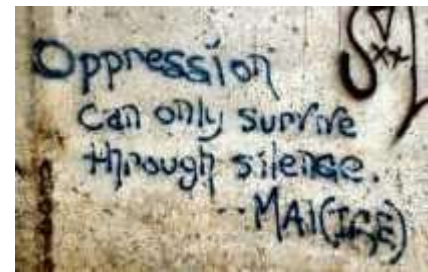
tooth and nail, only creating a dominant-submissive cultural relationship using aggression and scare-tactics.

## Commentary

It is clear from these thoroughly-encompassing aspects of cultural oppression—provided by Young, Hinson and Bradley and expanded on here—that cultural oppression exists despite its primarily unintentional nature. There is no single social group that purposely exhibits its dominance over another social group; however, that dominance is inevitably expressed culturally, socially and politically as the norm that excludes all that is foreign or different. The working class is still exploited by employers, who have differing views on fair wages and wage disparity; individuals who can not work, including for reasons such as lack of education and occupational experience, are marginalized, stigmatized and labelled; blue-collar workers, in general, are still excluded from workplace decisions that affect their careers and outside lives; the white, European population is still culturally dominant, while other populations fight for the opportunity to be culturally expressive and involved; and acts of violence, which perpetuate oppression and cultural subjugation, are still perpetuated against certain populations intentionally and unintentionally.



There is no doubt that these methods of oppression affect populations of color more than they do white, European populations. There are more individuals of color than white individuals in the working, blue-collar class, as well as in state and federal prisons; there are more people of color coming from low-income communities, translating into fewer people of color graduating secondary school and earning post-secondary degrees; and people of color are the least likely to be culturally represented but most likely to experience poverty and violence.



For a society in which “oppression” is an unacceptable topic of conversation, there sure is a lot to talk about.

## THE JEDI CORNER

# How to Help Advocate for Justice, Equity, Diversity, and Inclusion

Kimberly F. Sellers, ASA Justice, Equity, Diversity, and Inclusion (JEDI) Outreach Group Chair

The year 2020 proved to be a watershed period in our society, given the dual epidemics of health disparities and racism. It brought many underlying issues regarding justice, equity, diversity, and inclusion to the forefront, such that people could no longer turn a blind eye to them.

As you have since hopefully heard and/or read, the American Statistical Association has established the Justice, Equity, Diversity, and Inclusion (JEDI) Outreach Group to develop a diverse and inclusive environment that embraces all statisticians and data scientists, as well as to study and address statistical methods that can potentially introduce or exacerbate inequities and injustices in our society. Accordingly, we are committed to communication, programming, and professional development to advance and support a society that values all people.

Following through on this commitment, the JEDI Outreach Group hereby introduces one such initiative—The JEDI Corner. The JEDI Corner will be a regular component of *Amstat News*, where statisticians can write about and educate our community about various JEDI-related matters.

When most people hear commentary relating to JEDI, they typically focus on discussions such as those regarding race, gender, or sexual orientation; however, there are numerous other matters that likewise deserve attention in our community: persons with disabilities; ageism; and disparities ranging from financial/economic to discipline-specific (from theoretical to applied statistics) or with regard to work environment (academic, government, or industry).

Notice our potential span is enormous and, frankly, intimidating for one organization to address, which is why we need your help.

Issues of disparity and inequality are rampant in our society, so it is only natural they would exist within our discipline, as well. Thus, such matters affect everyone and, accordingly, everyone should care about at least some measure of advocacy.

We further hope for this regular segment in *Amstat News* to serve as an opportunity to educate our community about statistics research and developments

that potentially contribute to or impact JEDI-related matters. As such, we want to not only be mindful of the usual statistical implications of our developments but look beyond “the data” and consider the impact our work can have on the larger question for which the data was gathered, collected, and analyzed.

Consider also not only how our work affects “the data majority,” but how it affects the data lying on the periphery or perceived to be an outlier. Particularly when working in applications with real biomedical or societal implications, we cannot simply dismiss the data that appears as “different from the norm” or somehow misrepresents its contribution to convey a desired result more easily.

Notice that such statistical interests and contributions can span any number of topics. As a result, this is an exciting opportunity for any and all to contribute to this endeavor—from our vast array of sections and committees to interest groups. Whether or not you affiliate with any of these ASA subgroups, there are JEDI-related topics for discussion on which you can shed light for our community.

Any interested writers can submit a JEDI-related article for The JEDI Corner; submissions or inquiries regarding The JEDI Corner can be emailed to the outreach group’s communications team at [communicate@datascijedi.org](mailto:communicate@datascijedi.org) with “The JEDI Corner” in the subject line.

Contributions can be submitted at any time, and The JEDI Corner manager will be in touch with you regarding your inquiry or submission, any associated review, and publication logistics, as appropriate.

Submissions should be 600–1,200 words and, as with any publication, note the authors’ names and affiliations for proper crediting, along with a title for the contribution.

On behalf of the JEDI Outreach Group, thank you in advance for your contribution to educating our community and helping make our discipline and larger society a more just, equitable, diverse, and inclusive environment for all. ■





# Addressing **Implicit Bias** Among Women Statisticians and Data Scientists

*Kimberly F. Sellers, Emma K. T. Benn, Maria Garcia, Meghan Kellam*



*While implicit biases play a significant role and impact in gender relations in the statistics and data science communities, a variety of other factors differentiate women in these fields so that unconscious characterizations may produce implicit biases.*

The Kirwan Institute for the Study of Race and Ethnicity at the Ohio State University defines implicit bias as:

The attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner. These biases, which encompass both favorable and unfavorable assessments, are activated involuntarily and without an individual's awareness or intentional control. Residing deep in the subconscious, these biases are different from known biases that individuals may choose to conceal for the purposes of social and/or political correctness. Rather, implicit biases are not accessible through introspection. (<http://kirwaninstitute.osu.edu/research/understanding-implicit-bias/>)

While implicit biases play a significant role and impact in gender relations in the statistics and data science communities, a variety of other factors differentiate women in these fields so that unconscious characterizations may produce implicit biases. These factors motivate this article.

One such classification that potentially lends itself to introducing implicit bias is career choice (academia, government, or industry) and trajectory. Students, particularly at the graduate level, are naturally influenced by their advisors, other departmental faculty, or other academics with whom they interact. Those exchanges can influence a student's perception about career opportunities outside academia, perhaps implying such opportunities to be less appealing or fulfilling in some way. Because all individuals carry implicit biases with them, it is logical that academics would encourage and advise their students to pursue career options akin to their own.

That said, however, it is important for graduate students and early-stage professionals to identify mentors

and intern in various sectors to make informed decisions about career path options.

Government and industry internships can broaden perspectives regarding the broad range of career opportunities that are available with statistics and data science training. Another option is to hire a career counselor or related professional to help manage the job-seeking environment. Women who may temporarily leave the workforce (e.g., to care for their children) may experience doubts about resuming their careers, implicitly expecting to be judged harshly for taking time off from work. A career counselor can help navigate the nuances of retaking a professional life and preparing applications that account for the career gap.

Another area influenced or affected by implicit bias is divulging information about marital and/or familial status in the workplace. Women should be mindful about divulging such personal information in the workplace because doing so can introduce unexpected (and, sometimes, unwelcome) repercussions.

Women who are single without children are often perceived as either frivolous with their finances (because of the implications about their personal choices) or are believed to live for their work, with their personal lives being held with less regard than those having a family for which to care, often being called upon for extra or late-evening responsibilities in the workplace.

Meanwhile, women who are married and/or have children can be perceived as less loyal to their workplaces and even as capricious. While work environments have generally made great improvements through the years, cases are still reported of those where men can proudly display pictures of their spouses and children while women are leery of doing the same, because of the implicit biases that come with such displays. These simple gestures contain a disturbing undertone: By displaying family photos, men are viewed as stable and trustworthy, while women are viewed negatively as less devoted to their careers. Such biases create a heavy


burden on women, including striving to do their jobs even during significantly trying personal circumstances.

We cannot offer a general rubric on how to address the matter, but we have each been fortunate to have had positive experiences with understanding department or division heads who helped to address the situation at hand optimally. Nonetheless, it is important to be mindful of one's environment, recognize the important balancing act that exists, and determine which personal matters to share with colleagues or coworkers versus those that you prefer to keep private.

Another nuance associated with marital and familial status involves orientation in the LGBTQ community. Often, conversations in the heteronormative workplace dismiss marriages among LGBTQ colleagues or coworkers; this makes those in the LGBTQ community feel marginalized and invisible, and it should be rectified.

Racial/ethnic identity is most commonly associated with issues of implicit bias; this matter sparks much debate, often with confounding implicit biases due to gender status. Just as women have historically been viewed negatively both within and outside the statistics and data science communities, under-represented minorities face additional impediments due to both conscious and unconscious biases that others hold regarding people perceived as different from them.

The major difference between the previously discussed factors versus race/ethnicity and gender is that the other factors rely on information divulged by the individual, while one's gender or racial/ethnic makeup is more immediately identified through the senses. For example, while potential workplaces do not know (and cannot inquire about) marital status, they can immediately assess or infer gender and/or ethnicity when interacting with an individual (or even before) by seeing someone's complexion, hearing an accent, etc.

It is less clear how to address such issues when they confound with many other societal issues. In an effort to do so for the statistics and data science communities, we should create safe spaces for sharing experiences regarding both implicit and explicit biases. One such opportunity will occur through an implicit bias session at the 2017 Joint Statistical Meetings (scheduled for 2 p.m. on Monday, July 31, 2017), co-sponsored by the ASA Committee on Women in Statistics and the ASA Committee on Minorities in Statistics. 

## About the Authors

**Emma K. T. Benn** is a DrPH and assistant professor in the Center for Biostatistics and Department of Population Health Science and Policy at the Icahn School of Medicine at Mount Sinai (ISMMS). She is also co-director of the Master of Science in Biostatistics Program at ISMMS and co-chair of the Eastern North American Region of the International Biometric Society Fostering Diversity in Biostatistics Workshop. Her recent methodological interests consist of applying novel statistical approaches to either correct for or draw attention to selectivity bias in observational studies and clinical trials.

**Maria Garcia** has a PhD in applied mathematics and experience working in academia and government. She is a research mathematical statistician in the Missing Data Methods Research group in the Center for Statistical Research and Methodology at the U.S. Census Bureau. Her research focuses on developing methods for statistical data editing, imputation, model-based imputation, and multiple imputation.

**Meghan Kellam** holds a degree in psychology from Northeastern University. She worked in special education for three years before transitioning into the biopharmaceutical industry, where she has worked for more than 15 years. She has held a number of positions in service providers, consulting, and biopharmaceutical companies. Her current position is on the leadership team in the Biometrics department at Biogen, where she leads a team of program data and analytics leads responsible for programs in the immunology and neurology portfolio.

**Kimberly Sellers**, PhD is an associate professor of mathematics and statistics, specializing in statistics, at Georgetown University in Washington, DC. She held previous faculty positions at Carnegie Mellon University as a visiting assistant professor of statistics and the University of Pennsylvania School of Medicine as an assistant professor of biostatistics and senior scholar in the Center for Clinical Epidemiology and Biostatistics (CCEB). Her areas of interest and expertise are in generalized statistical methods involving count data that contain data dispersion, and in image analysis techniques, particularly low-level analyses. She is the 2017–2018 chairperson of the ASA Committee on Women in Statistics.

This report is intended to inform parties of ongoing research and to encourage discussion. The views expressed are those of the authors and not necessarily of the U.S. Census Bureau.

## Conversation

# A conversation about implicit bias

Amanda L. Golbeck<sup>a,\*</sup>, Arlene Ash<sup>b</sup>, Mary Gray<sup>c</sup>, Marcia Gumpertz<sup>d</sup>, Nicholas P. Jewell<sup>e</sup>,  
 Jon R. Kettenring<sup>f</sup>, Judith D. Singer<sup>g</sup> and Yulia R. Gel<sup>h</sup>

<sup>a</sup>Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR, USA

<sup>b</sup>Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA

<sup>c</sup>Department of Mathematics and Statistics, American University, Washington, DC, USA

<sup>d</sup>Office of the Provost, North Carolina State University, Raleigh, NC, USA

<sup>e</sup>Department of Biostatistics, School of Public Health, University of California – Berkeley, Berkeley, CA, USA

<sup>f</sup>Research Institute for Scientists Emeriti, Drew University, Madison, NJ, USA

<sup>g</sup>Office of the Provost, Harvard University, Cambridge, MA, USA

<sup>h</sup>Department of Mathematical Sciences, University of Texas – Dallas, Richardson, TX, USA

**Abstract.** Explicit bias reflects our perceptions at a conscious level. In contrast, implicit bias is unintentional and operates at a level below our conscious awareness. Implicit stereotypes shaping implicit biases are widely studied in criminal justice, medicine, CEO selection at Fortune 500 companies, etc. However, the problem of unconscious bias remains. E.g., while women constitute an increasing proportion of all STEM undergraduates, they still make up only a small proportion of faculty members at research universities, and they are substantially under-represented in organizational leadership and as recipients of professional awards and prizes. Can we afford to have unintentional perceptions continue to hinder the success and advancement of women and other underrepresented groups? Can we afford to continue to underuse human capital in science? This session at the 2015 Joint Statistical Meetings (JSM) aimed to illuminate what statisticians need to know and do to break the glass ceiling of implicit bias.

**Keywords:** Stereotypes, leadership, awards, women, minorities

## 1. Introduction

**Golbeck:** Why did I want to organize a session at the JSM on “Implicit Bias: What Statisticians Need to Know and Do”? I will answer this question with three stories.

One relates to a video. In the late 1990’s, one of my responsibilities was to direct a university-wide mentoring program. I hired a diversity trainer who brought in a video, *The Color of Fear* [33], which captured deep conversations about race among men gathered for a weekend retreat. We watched as the men discovered racist attitudes they didn’t know they had. It was powerful.

A second story relates to an interview. A few years ago in relationship to a microhistory project, I interviewed a former UC-Berkeley law school dean, Sanford Kadish. The background is that, in 1968 at Berkeley, only 3% of tenure ladder faculty members were women, and 44% of departments had gone over 50 years without having even one tenured woman. I asked Kadish why it was that Berkeley first took up the issue of equal employment opportunity for women in 1968. He said, “Academic senate policy committee members were so busy thinking about discrimination against ethnic minorities that they just hadn’t ever given thought to discrimination against women” [11].

A third story relates to the *JSM Registration Guide*. When I received my guide in 2012, I noticed there were no pictures of women in it [10]. In 2013, I noticed that, of the 15 JSM keynote speakers, only one – the American Statistical Association (ASA) president – was a

\*Corresponding author: Amanda L. Golbeck, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, 4301 West Markham, #820, Little Rock, AR 72201, USA. Tel.: +1 501 526 6611; E-mail: agolbeck@uams.edu.



Fig. 1. The Panel (l to r): Jon Kettenring, Marcia Gumpertz, Arlene Ash, Judy Singer, Mary Gray, Nick Jewell, Amanda Golbeck (moderator). Photo credit: Yulia R. Gel.

woman. Yet, about one third of ASA members were women.

When I brought the pictures issue to the attention of ASA leaders in 2012, they invited me to attend a national meeting of the AWARDS project in which they were participating. “AWARDS” is an acronym for Advancing Ways of Awarding Recognition in Disciplinary Societies. The National Science Foundation (NSF) funded this project in collaboration with the Association for Women in Science (AWIS). The question for the ASA was, if one third of members are women, then why don’t one third of the ASA awards for research go to women? AWARDS focused on the role that implicit bias plays in the underrepresentation of women among research award winners [2].

AWARDS put a name to what I saw in the video, interview, and JSM Registration Guides. It was implicit bias.

Unfortunately, implicit bias problems for women and other underrepresented groups have been around for a long time. For example, in 1984 the late UC-Berkeley Statistics Professor Elizabeth L. Scott looked at the percentages of NSF fellowships to women. She noticed that women won few such awards and wrote: “I wonder what is happening. There must be many different committees (panels) ranking the applicants in the different fields. Why is it that women are still coming out so poorly in the percentage of awards? Have there been some studies made that help to explain . . . these striking differences in the probability of obtaining an NSF Fellowship, once you apply”? [28].

Yulia Gel (University of Texas-Dallas) and I (University of Arkansas for Medical Sciences) organized the JSM panel on implicit bias because we want our statistics organizations to be big tents for statisticians of all areas of practice, both genders, and all races and ethnicities. We want the same for our workplaces.

We invited to the panel ASA Fellows Arlene Ash (University of Massachusetts Medical School), Mary Gray (American University), Marcia Gumpertz (North

Carolina State University), Nick Jewell (University of California-Berkeley), Jon Kettenring (Drew University), and Judy Singer (Harvard University) (Fig. 1). These are statistician-leaders who have thought, and continue to think, deeply about implicit bias issues. Several are statistician-administrators who work out of provost’s offices and specialize in issues of equity in employment. Sponsors of the session included the Caucus for Women in Statistics (the primary sponsor), Joint Committee on Women in the Mathematical Sciences, International Chinese Statistical Association, International Indian Statistical Association, Statistics Without Borders, ASA Section on Statistical Consulting, and ASA Scientific and Public Affairs Advisory Committee.

The panel organized around five primary questions. These provided the framework for the session and the report that follows.

## 2. What is the difference between implicit and explicit bias, and why do we care about bias in the profession?

**Golbeck:** When I think about bias, I recall a paper that I discovered early in my career, “Bias in Analytic Research”, where the author created a useful catalog of 35 biases that can distort the design, execution, analysis, or interpretation of research [26]. Implicit bias isn’t in the catalog; there only are a couple of biases where you can surmise a connection with implicit bias. An example is: “Expectation bias. Observers may systematically err in measuring and recording observations so that they concur with prior expectations”.

So we can’t extract an easy definition of implicit bias from our own literature. Jon, you have examined the broader literature. Can you please tell us what you’ve found?

**Kettenring:** I’d first like to add to my credentials. I manage a group of about ten people that has been in ex-



istence for about 35 years. Our group has had only one woman in 35 years, and she is now retired. So perhaps I should stop here!

Just to make sure we are all on the same footing, let me start with some basic ideas of implicit bias. One of the things I ran across in preparing these remarks is a large study from Ohio State University by the Kirwan Institute for the Study of Race and Ethnicity [24]. My early comments will be based on what I found there. If you don't know it – and I did not – there is a rather large literature on the topic of implicit bias.

According to the Kirwan report, implicit bias refers to the attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner. Here the key word is unconscious. Some of the other things they emphasize that go beyond the unconscious and unintended are that the biases we are talking about are robust and pervasive, and we all – you and me – are susceptible to them.

On the other hand, at the other extreme, we have the explicit biases. These involve a higher degree of awareness. They may involve willful ignorance and blatant discrimination. These sorts of biases can be consciously detected and controlled, and some people claim they can be measured using various instruments.

As just a slight footnote, in thinking about these differences, the discrete separation between implicit and explicit may be more for our convenience than reality. Different ones of us would probably draw the line in different places.

Coming back to implicit bias, there has been a lot of research and debate on this topic for at least 30 years. Sometimes instead of using the term “implicit”, you'll find people referring to it as unconscious bias, or some other similar term. Another label that is quite common is “automatic”, as in instinctive or a gut reaction. So a lot of the implicit biases involve various types of stereotypes that we all no doubt have to some degree at least.

Some people have pointed to a connection in a modern book that many of you have probably read and I am presently reading. This is the book by the psychologist Daniel Kahneman, *Thinking, Fast and Slow* [18]. The thinking fast part of it ties in very nicely to the issue of implicit bias. The most common context for the various studies that I have seen involves implicit bias for women and minorities. But if you think about it for a minute there are lots of other potential applications that are perhaps less noteworthy: rich versus poor, old versus young, and so on. The Kirwan report that I referred to earlier talks about implicit biases concerning

racial disparities that arise in our educational system, the judicial system, and the health system. In a recent issue of *Science* [29], I noticed a very interesting article that talked about cognitive bias on the part of forensic scientists. My take on it is that this is just another example of implicit bias.

Why do we care about implicit bias in the profession? I think the reasons are pretty obvious. I will mention a few. Implicit biases impact how we deal with one another, and the decisions that we make—decisions about hiring, pay, admissions, promotions, awards, committee assignments ... much of what we do as professional statisticians. The bottom line is that these implicit biases can result in poor decision-making, and consequentially generate all kinds of inefficiencies and distortions that we should not tolerate, and we really can't afford to have them.

**Golbeck:** Arlene, thinking as a statistician, you've noticed a possible relationship between implicit bias and Bayesian logic. Is implicit bias just Bayesian logic and, if so, what can we do about it?

**Ash:** Our society has similar numbers of men and women but more male scientists than female scientists. So the probability of being a scientist is higher for a man than for a woman, and – especially in fields where there are very few women – it's “natural” (Bayesian) to think “scientist  $\rightarrow$  male”. OK, scientists today are more likely to be men, but what does that tell you about whether, *between comparably credentialed people* applying for a job, one male and the other female, one has a higher probability of being a good scientist? Or whether the inquisitive little girl is more or less capable of becoming a great scientist as the inquisitive little boy? Understanding Bayes' theorem can help explain why people might bring automatic prejudices to situations in which they are unhelpful. I liked the following advice that emerged during the e-mail exchanges that preceded this panel: “If you are not confident about the suitability of an apparently well-qualified individual candidate, ask yourself, ‘Why’?”

**Golbeck:** Marcia, we're hearing a lot about microaggressions today in the field of psychology. What are microaggressions, and speaking generally, what can be their cumulative effects on individuals in the profession who experience them?

**Gumpertz:** Microaggressions are the manifestations of unconscious bias. People have different ideas, as Jon said, about where the line is drawn between implicit and explicit bias, but in any case microaggressions can be very tiny things. The word sounds aggressive, but microaggressions aren't necessarily hos-

tile or confrontational. Often they are just small slights, or little things that people say, and people aren't really aware that they're doing them or saying them. You can be in a situation where somebody does something and you aren't really aware that it's happened, unless you're the recipient of it. Microaggressions can be hard for the recipient to talk about, because people might think they are just being too sensitive if they react to these kinds of things.

Let me tell you about a few of the kinds of things I'm talking about. This is a familiar scenario: A woman is in a meeting and says something, and somebody else interrupts or somebody right behind her says something else, and the conversation goes on and nobody ever acknowledges that she said anything. Or she says something and nobody picks up on it, and a little while later somebody else says essentially the same thing – maybe a man says it – and then everybody realizes this is a wonderful idea. Another kind of microaggression occurs when someone comes into a room and looks around and says hello to everybody except the single person of color in the room. And another one – this is one a department head told me about – they had done a department climate survey, and in the survey somebody said there was a male faculty member in the department who always put his arm around the women faculty and staff, and the department head didn't know what to do about it.

The recipient of a microaggression often doesn't want to say anything because others might just react by saying, "You're being too sensitive, you should just ignore it and let it go by". That kind of reaction can also be viewed as a kind of microaggression, indicating that the confidant is not really listening to the recipient of the microaggression and doesn't believe their interpretation.

Why are these important? Well there are some groups of people who experience this kind of thing in many interactions, even on a daily basis. The impact builds up, so they become sensitized to it. Their reaction may be to say to themselves "I don't really want to deal with these people anymore" and to disengage. That is really destructive for the group and the organization, and it is destructive for the success of the person who is disengaging.

One important thing to me about microaggressions is to learn to recognize them. It is important to figure out ways to interrupt them if you see them happening when they are not happening to you. That's one thing. But really the first thing, and the big thing for me in my journey along this path, has been to recognize when

I am perpetrating some kind of microaggression, and just really seeing it. There are things that we all do, and it's not easy to realize you are actually doing these things. But that's the first step in changing how the department climate is, for example, in your organization.

**Golbeck:** So you are talking about taking things that are more in your subconscious and bringing them more into your conscious awareness.

**Gumperz:** Right.

**Golbeck:** Marcia has pointed out that microaggressions, which result from unconscious biases, can also be destructive for individuals. Nick, which do you think is more important in affecting the growth of women or underrepresented minorities in an organization: Is it implicit bias or explicit bias?

**Jewell:** At first glance, I think most of us would agree that the consequences may be more extreme in cases of explicit bias, in terms of sexual harassment or overt discrimination, people losing their jobs, being damaged psychologically. But I think implicit biases are actually more difficult to deal with, and their consequences may be much more pervasive and destructive in the long run. That's worth thinking about, because we all feel strongly about blatant discrimination or illegal acts between men and women or people of different races. We believe laws should be enforced. But most people have a much weaker reaction to cases of implicit biases, as just indicated, when people tend to say: "You're being too sensitive, let it go, it doesn't matter". But there are consequences in the long run.

President Obama recently said in a renewed US conversation about race that the confederate flag should be taken down; he was urging its removal from the state capital in South Carolina. He went on to say that, while the confederate flag can be taken down, this doesn't mean it is gone from the heart. Obama's comments made it clear African-Americans generally appear to be much less happy and more dismayed about persistent implicit biases that they face every day than they are about people who obviously exhibit explicit biases towards them. While this might strike you as odd, it reflects how pervasive these implicit biases are and how deeply they are felt.

On a lighter note, you might wish to listen to a recent video interview (available online at <https://www.youtube.com/watch?v=J-VzAw2HQq4>) between Kristen Stewart and Jessie Eisenberg, young actors in a current movie that Hollywood is advertising. They did a parody where they were going to interview each other to promote the movie. They flipped the cards and gave the 'woman' questions to Kristen to ask Jessie, and

the ‘men’ questions to Jessie to ask Kristin. Jessie asks Kristen: “Tell me about how you were in high school, were you an actor, were you a class clown”? She answers that question very easily. Kristin turns to Jessie and says: “Are you in a relationship with someone right now? Are you pregnant”? He reacts: “That’s a little personal, isn’t it”? Then Jesse asks Kristin: “What is your favorite sports team”? She says “the Lakers”. Then Kristin turns to Jesse and says: “Who designed your shirt”? He says: “I don’t know, Levi? Who Cares”? As the interview proceeds, things get more heated, and finally Kristin turns to Jesse and says: “Now you know what it feels like to be a woman in Hollywood”. This dialog prompts us to ask: What ‘question cards’ are you using when you look at someone’s resume? What questions are you asking when you invite a speaker? What card deck are you using when you are voting for someone to get on a board? I think you’ll be surprised at how significantly implicit biases are affecting the questions you are thinking of depending on the ethnicity, gender, or race of the person you are thinking about.

**Golbeck:** So far we’ve heard some of the everyday perspectives on implicit bias, but there may also be legal perspectives. Mary, you are both a statistician and an attorney. What would you say are the legal perspectives on implicit bias?

**Gray:** I’m not convinced there is a difference between implicit bias and explicit bias – so I probably don’t belong on this panel! It seems to me the effect is the same no matter what the kind of bias. It is not good.

What about legal implications, then? There is a concept in law that is somewhat similar to implicit bias, and that’s the issue of apparent discrimination in “disparate treatment and disparate impact”, both of which are illegal [15].

Disparate treatment is: I’m not hiring or promoting you because you are a woman, or because of your race or some other reason.

Then there’s disparate impact. The most recent US Supreme Court case that had to do with disparate impact was a United Parcel Service (UPS) case involving pregnancy rather than sex per se. The UPS had a rule that you couldn’t be a driver if you couldn’t lift 50-pound weights, because you had to deliver the packages. As a result, when a woman was pregnant, and her doctor told her she couldn’t lift 50-pound weights (incidentally, think about how many of you can lift 50-pound weights, particularly up over your head), she was put on unpaid leave. But when the case came to court, it turned out the heaviest package the USP

ever had on these routes was 20 pounds. Therefore, the court decided this was not a reason for putting the pregnant woman on unpaid leave. Going further back, there is a Supreme Court case that has to do with medical benefits, before we all had Affordable Health Care like we do now. What it said was: Not covering pregnancy benefits was OK, because pregnant men weren’t covered either; this was discrimination on the basis of pregnancy, not discrimination on the basis of sex. We went through a round with pensions where there was discrimination on the basis of longevity rather than discrimination on the basis of sex. But they are all discrimination, and they are all bias, and the origin is the same.

Although we talk about disparate treatment and disparate impact, the impact is actually the same. It has a bad impact. So, the clarity that we get in the law is that both of these are illegal. This is the clarity that we need to understand in talking about implicit and explicit bias. The reason I said I don’t see the difference is that I view implicit bias as a concept that someone came up with so you can show lots of slide shows to people about implicit bias, and then go away saying we have done something about bias. I sat through a number of these, one of which was with the ASA in how we select people for awards. There was some improvement as a result, but there is a tendency to slide back.

In fact, the tendency to slide back is something we all need to be cautious about. I was in a panel this past week in which we were talking about the percentage of PhD’s in mathematics that go to women. It turns out that finally in 1988 we got back to where we were in 1930. It had gone down to 6% of PhD’s in math being women at the time I got my PhD, but it got back up to the low 20’s by 1988, and it’s gone up a little bit more now. Why? Because they have now included statistics in with math, and proportionately lots more women get PhD’s in statistics than in math.

So, implicit or explicit, we need to worry about it.

### 3. What are some concrete examples where you believe you’ve seen the effects of implicit bias on evaluations of women and underrepresented minorities?

**Golbeck:** Next let’s turn our attention to some

concrete examples of implicit bias. Following the AWARDS meeting (that I mentioned earlier), I examined the ASA awards in more detail. From 2001 through 2012, the ASA gave seven types of research awards, and only 12% went to women, whereas 33% of ASA members were women. We do pretty well at selecting women to be presidents of the ASA, and fairly well at giving women awards for teaching and service, but we don't recognize enough women for their research [12]. There were no women at all among any of the Deming Lecturer Award winners. Likewise, no women won the Dixon Award for Excellence in Statistical Consulting. The research award bestowed most frequently on women was the Outstanding Statistical Application Award. But even there, only 16% of winners were women. It seemed to me that women did better at winning awards where the judgment was based on a particular paper or group of papers, as opposed to lifetime achievements.

Judy, what are some other examples of implicit bias on evaluations of women and underrepresented minorities?

**Singer:** To set the context, in addition to being a professor of statistics at Harvard University's Graduate School of Education, I am also the Senior Vice Provost for Faculty Development and Diversity for the entire university. In this role – which I've had for seven years – I oversee faculty affairs and academic appointments. Of particular importance to today's discussion, I participate in every tenure review. At Harvard, we solicit confidential letters about each potential tenured appointment from outside experts and also from the department or school's tenured faculty. I read all of these letters.

Let me give you a sense of what I sometimes see by sharing examples (each with identifying details changed to protect confidentiality). I keep these in a file labeled "outrageous letters".

- "I have serious concerns about the process in this search. The identities of the two finalists make it clear that the search was rigged from the start. How else could both be female"? As you might imagine, there are many searches in that department historically in which the two finalists have both been male. But to this senior faculty member, those could not have been rigged; only the search with two *female* finalists could be rigged.
- "The following anecdote makes me a bit unsure about our judgment. The last African-American candidate we interviewed was X. The consensus then was that she did a very poor job field-

ing questions and we wondered about her quality of mind. The paper she presented has since been published in [the top journal in the field]. Clearly she has been very successful at [University of Y] and in retrospect, I wish we had made her an offer". So they then tried to recruit her. And what did she do? She turned them down. Had she been appointed in the first place, they might have been successful at catching her at an earlier career stage.

- "She balances work and life in a way that detracts from her career". When you read confidential letters, you often find that women are nice, have children, and balance work and career. I'm confident that most of the people who write these things truly believe they're being helpful, but I'm here to say that they are not!
- "And what's more remarkable is that she did all of this while having three children". This sentence was particularly jarring because it's from a letter for a search in which we were recruiting a husband and wife, who also happen to write many papers together. Some external letter writers were asked to comment on only one of the candidates and a few were asked to comment on both. This particular letter writer made it very clear up front that when he was going to write about their "joint work" he would use the same paragraph, and when he was going to write about their separate work he would use different paragraphs. In the parallel paragraph about the husband, that included much of the same text, the letter writer didn't mention that "what's more remarkable is that *he* did all of this while having three children" – even though the wife's children are indeed also the husband's. If you were to draw conclusions about faculty demographics from reading confidential letters, you'd conclude that only women have children (as you rarely read about family matters in letters about men).
- "Her speech accepting the X award was the high note of the ceremony... eclectic, passionate, and well delivered; as well, her kimono was spectacular". Off-hand remarks like this serve to remind readers that the candidate is a woman and what she's wearing is important. On this point, here's another anecdote: I have two colleagues – one male, one female – who were invited to a major White House conference. We debriefed them about what happened at the White House and what kinds of questions they were asked when

they were back on campus. The man was asked about which part of his excellent research he focused on. And what was the woman asked? What did she wear!

These examples are not the norm, but comments like these appear all too often. Implicit bias influences what people write and what people say more than most of us care to know.

**Golbeck:** There are so many areas in which we see these kinds of behaviors in the workplace. Marcia, Judy's examples had to do with faculty. I'm wondering if you could please provide some examples that have to do with students. Especially, what are the effects of implicit bias on academic advising?

**Gumpertz:** This topic is very important. There is a gap. I think that a good percentage of doctoral degrees in statistics are going to women. But then there is a drop in percentage of statistics faculty positions held by women. What I'm wondering about is the advising that the graduate students get about how to select their advisors and dissertation topics. In statistics, you can do many different kinds of dissertations, and they'll be suitable to prepare for a wide range of statistics career paths. I really have a question about whether women and minority students are being unconsciously steered toward nonacademic career options.

On the topic of selecting a major professor, an article came out recently looking at women in the biological sciences [30]. The authors examined the gender distribution of graduate students and postdocs who worked with faculty who had received prestigious awards. What they found was that the labs that were headed by prestigious male faculty trained a much smaller fraction of women than were in the graduate student population. In the biological sciences more than 50% of the PhDs go to women, but men with prestigious awards did not train nearly 50% women in their labs. In contrast, the women with prestigious awards did train a high fraction of women. This is important because a large number of people who go on to academic careers in these disciplines come out of these labs.

**Golbeck:** So we have heard about some examples of implicit bias that involve faculty and students. Nick, what can you tell us about the effects of implicit bias on selection of academic administrators?

**Jewell:** Since leaders set many policies for the rest of the community, implicit bias in the selection of administrators is a serious issue. I'd like to talk about my experience with leadership positions in higher education, where the situation may be much worse than in the private sector. Here I will present two symbolic

cases concerning the appointment of university chancellors/presidents.

In the first case, a male told a story of his appointment as president of a very prominent public university. After he had been in his new position for several months, a regent who had served on the selection committee took him to lunch. This regent told the man that he had liked him instantly because he "looked like a president – you're tall – not like that short stubby guy [the other candidate]". So while it is generally understood that being tall as a man tends to make him more successful and attractive, it apparently sometimes is also the key to getting a university leadership position! One can only wonder how this regent was making other decisions, perhaps even more implicitly.

The second case has to do with the selection of a university chancellor – also at a leading public university. Two candidates, one male and one female, were asked a question about how they would handle a decision in the face of forcefully held and diametrically opposed opinions. A male regent indicated that they should select the man because his answer was to choose one position decisively, whereas the woman had suggested a more nuanced and cooperative approach: "She's going to compromise. You have to be decisive". The regent's response illustrates that the 'woman' answer can often be discredited by men because they simply "don't get it". We all suffer from this kind of implicit bias. We all often "don't get it".

These two cases refer to the choice of leaders of two major universities, people who are ultimately responsible for setting policies about promoting equity. These examples reflect the extraordinary uphill battle we have in rectifying the lack of women in leadership. Some countries have very family-friendly policies for women. Scandinavian countries have long had favorable policies for women, and they are much better places to work if you are a woman; but unfortunately they have no better record on having women in levels of leadership, like prime ministers and leaders of big companies, than Americans do. So there remain some real issues to deal with regarding women in leadership.

**Golbeck:** You noted an implicit bias of how men should look to appear presidential. The number of women in upper levels of leadership is unbelievably small [13]. Do we have a conception of how women should look to appear presidential?

**Jewell:** Well, we haven't had a woman president in the US yet and Berkeley has never had a woman chancellor, so we don't really have a conception. Presumably Margaret Thatcher is what a woman prime minister "should" look like!

**Golbeck:** Mary, what are some specific examples of the effects of implicit bias on the advancement of science?

**Gray:** Statisticians and mathematicians are key players in a lot of studies in science. A lot of times bias gets reflected in things that we get parts of. For example, it used to be that clinical trials only had males as the subjects, and nobody thought anything about this. You can obviously see that there is likely to be a different result if you had both men and women. It was the women scientists at NIH who agitated until the rules got changed to include women in clinical trials, and this change to include women has extended to a lot of other research that's being done.

I have a number of other examples just from the class that I teach on survey sampling. We did a project last fall during the national elections that involved whether or not people could vote, because they either had or didn't have voter ID's in the state of Virginia. One of my students said: "But you know, everyone has an ID, how can you fly without an ID"? Well, American University is the kind of institution where most people have the resources for air travel, and the student's concepts were limited to that reality. This isn't sex bias, but it is definitely a kind of bias.

In another semester, my students were going to do a study to determine whether foods that were healthy cost more than foods that weren't, and therefore people were not eating properly. They did surveys in various parts of Washington DC. Their results showed, contrary to their assumptions, that people were paying much more for food that wasn't healthy than they were for the food that was. Why did this happen? Because the young woman who organized the survey was a vegetarian, and she totally neglected the effect of the price of meat in the total cost of groceries.

So you can see there are all kinds of things that can be influencing science more generally, not just whether you get a job, whether your pay is equal, or whether you get to be president of a university.

**Golbeck:** Arlene, it has been shown that it isn't just people who have implicit biases. Even clever data-mining algorithms can inherit implicit bias [19]. Can we use computerized algorithms to reverse-engineer bias?

**Ash:** We had a lively pre-panel exchange, finding and sharing many interesting articles. A recent one in the *Washington Post* described how researchers at Carnegie Mellon University studied Google's algorithm for placing job ads in front of people whose search history suggests that they might be interested

in them [6]. Of course, Google won't tell you how its job ad-placing program *works*, but the Carnegie Mellon "Ad Fisher" team devised a test to see how it *performed* in placing these ads.

The program created thousands of different accounts. Each simulated the web-search behavior of one "simulated person" looking for a job. The machine created the same search history for simulated individuals who differed only by sex. The results showed over 1500 "men" but only 300 "women" received ads for high paying, high prestige jobs.

Of course, a sophisticated algorithm would seek to place ads based on the fraction of similar placements that had been "successful" in the past. Maybe women have less often clicked on such ads; or maybe employers advertising such positions have rarely offered such jobs to women. Whatever the underlying mechanism, such algorithms instantiate, perpetuate, and potentially amplify our society's prejudices. Google has a "bias busters" internal team. I hope that they can figure out how to deal with this algorithmic bias.

We can all learn from the example of the NIH Common Fund's High-Risk, High-Reward Program, which found that a search for "risk-takers" nets fewer women than one seeking "innovation" and "creativity" [5]. When looking for leaders, let's avoid sending signals that artificially limit the pool of potential talent.

#### 4. What has been the effect of all the attention to so-called 'implicit bias'?

**Golbeck:** There has been a lot of attention to these issues both in organizations and the media. An example is the popular book, *Blink: The Power of Thinking Without Thinking* [9]. One effect is more attention is being given to the neuroscience of implicit bias. I'm talking about the fMRI-based studies that are the subject of another popular book, *How Your Unconscious Mind Rules Your Behavior* [20]. fMRI stands for functional magnetic resonance imaging. This technology is a research tool that psychologists are using to combine MRI and behavioral data. This powerful combination is allowing us to learn more about some extremely difficult subjects like prejudice and stereotyping.

Now let's turn to some of the effects of implicit bias that our panel would like to talk about. Mary, in your experience, what has been the effect?

**Gray:** I don't actually think there has been very much effect. I think it's been distracting. You are going to hear about how things have improved in various

organizations. But I've been on panels and committees over many years, and there is always some reason why results of studies don't get implemented. Either there is not enough money, or you didn't consider enough variables, or you didn't have a good sample, or something else was wrong.

The same is true with the implicit bias, where there is a lot of research about what so-called implicit bias is supposed to do. But my feeling is it is just a distraction, and it keeps people from getting down to the nitty gritty of what is the actual impact of the bias. It's back to the notion that we have in the law: What's been the effect of this, and what would change the effect? Just talking about it is not good enough.

**Golbeck:** We are seeing a small but growing body of research literature now that illuminates some of the effects of implicit bias. Marcia, how are you using some of this literature in your institution to try to change the effects of implicit bias on faculty searches?

**Gumpertz:** Many universities across the country, in particular many that have had NSF ADVANCE grants, have started training faculty search committees on diversity and the effect of implicit bias. I want to tell you about two studies that we use at North Carolina State University in facilitated discussions with faculty, so people on search committees can start to think about how this might come into play in faculty searches.

One of the studies was done about 12 years ago [4]. The researchers looked at 1,300 applications for jobs in Chicago and Boston. These were jobs in sales and clerical service-type occupations. The researchers made up a bank of resumes based on real resumes and submitted four resumes for each job ad: They assigned an African-American-sounding name to two of the resumes, and a white-sounding name to the other two. They submitted these resumes, and then they counted how many call-backs each resume got inviting them for an interview. For the white-sounding names, they had to send about 10 resumes before they got a call-back; for the African-American-sounding names, they had to send 15 resumes.

The other study was done more recently and in an academic setting [21]. You might think things are changing enough that this kind of thing isn't happening any more, but it is. Also, you might think that university faculty wouldn't fall into this kind of bias in looking at resumes, but they do. In this study there wasn't actually a job. The researchers made up one resume from an undergraduate student aspiring to go to graduate school. The student was applying for a supposed job as a lab manager. The researchers sent the resume

to 127 faculty members in the sciences and asked them for feedback on the resume. Their cover story was that they were developing a mentoring program, and they wanted feedback to help them develop the program. For half of the requests for feedback they put the name John on the resume, and for the other half they put the name Jennifer. Same resume. For every measure of faculty feedback – hireability (would you hire this person), competence (how competent do you think this person is), and mentoring (would you mentor the person) – the male candidate got higher scores than the female candidate. When asked how much they would be willing to pay the person, the faculty respondents said they would offer \$5,000 more to the male candidate than to the female.

**Golbeck:** So we have research that has found evidence for effects of implicit biases in job searches, and this research can be used to try to make search committees more aware. Judy, what other decision-making bodies in our organizations need to be made aware of the potential effects of implicit biases?

**Singer:** I don't use the phrase "search committee training". Most Harvard faculty members don't want to be "trained", nor do they think they need to be "trained". We've taken a somewhat different tack. Instead of presenting topics like implicit bias in a "training" context, we present them in the context of "here are some interesting things about yourself about which you might not be aware".

I am privileged to have Professor Mahzarin Banaji, one of the developers of the concept of implicit bias, as a close colleague. She also serves as the senior advisor to the Harvard University Dean of the Faculty of Arts and Sciences on these issues. We recommend every faculty member take an Implicit Association Test (IAT). I highly recommend that you visit the "Project Implicit" web site [3]. The IATs are little timed tests. The first one takes about 10 minutes. They're designed to help you learn, in the privacy of your own home or office, whether you might actually hold some implicit biases. They cover a wide range of topics, including race, gender, sexuality, and religion. Try taking one. If the first one doesn't reveal a bias, take another one. I can guarantee you'll discover some biases you don't think you have. We're all biased. We all think we're fair, but we're all biased. And we've found that faculty members are fascinated by the cognitive dissonance that taking the IAT creates.

In preparation for this session, I took an IAT earlier today so I'd have the feeling fresh in my mind. I've taken a lot of these over the years, and I'm sup-

posed to be the person at Harvard who has overcome this! Sadly, I was two standard deviations below the mean (but I won't tell you on which dimension). It's a little depressing when you find you're confronted with something about yourself that you can't believe is true. Most of us don't want to be prejudiced. But when you find yourself confronting behavior at odds with your beliefs, it gets you thinking that maybe your judgments can be clouded. With that, faculty become interested in learning about better ways of conducting faculty searches, about being more explicit and structured about criteria. If you could blind search committee members to the demographics of the candidates – and obviously you can't, but if you could – there's lots of evidence that suggests that it would make a difference. So I think there is a lot of work we can do without using the word "training".

We've also found that faculty aren't necessarily skilled at some of the organizational behavior skills needed to chair a search committee (or to be a committee member). So we've developed a relatively short 18-page guide on *Best Practices for Conducting Faculty Searches* [22]. It includes a special pullout section on implicit bias, but it also devotes a lot of attention to how to run (and participate in) effective meetings. I knew I was onto something when a member of the Harvard Corporation (who is the CEO of a major company) said, "This is really useful! I spend my life running meetings and I haven't thought about some of these things". So we teach colleagues not just about implicit bias but also how to run an effective meeting, how to listen to different voices, and how to prevent the grand poobah from dominating the discussion with pronouncements like "I'm sure you all agree that X should get the job". Implicit bias is a major issue but it's not the only area in which we need to improve. Many faculty members are quite intimidated by the process of chairing a meeting of their colleagues and they welcome help about these very practical matters.

The second thing I would say is that I think leadership makes an enormous difference. I'm fortunate that the current president of Harvard University is Drew Faust. She has been president for the last eight years. My prior president was Larry Summers, who famously made some comments about women in science. It was a very hot topic on the Harvard campus, which catalyzed energy on the campus and led, in part, to the creation of my office. That was a very concrete leadership response.

Deans can also make a huge difference. I've seen what happens when a dean says to a department chair

during a tenure review: "Is there any possibility that your judgment might be clouded by implicit bias"? All of a sudden implicit bias gets talked about back at the department. So leadership makes an enormous difference, and if you can talk to your department chairs, deans, provosts, and presidents, and raise this issue, perhaps using sample letters like the ones I read earlier, you might start to convince people that implicit bias may cloud people's judgments.

**Golbeck:** Arlene, you work in a medical school, and medical schools can be different environments. What's happening in your medical school in relationship to implicit bias?

**Ash:** I'd like to further address the tension that Mary discussed: Making people aware of their implicit biases is only useful if it leads to more diversity in places like medical schools and, more generally, developing every individual's potential. Our school does encourage students to take the implicit association tests, and some people push back: "That doesn't mean that I'm prejudiced". But the discussions raise important issues that our students might not normally think about. And many – even those who don't think that they are prejudiced – say they think they will behave differently because of this training. And maybe they will.

In talking to the woman who is in charge of diversity efforts on our campus, and in interacting with this panel, I came away with several thoughts. One: where you sit, who you are, and what job responsibilities you have shape your ability to actually advance diversity. Our school is concerned that less than 2% of our faculty members are African-American. However, only 2% of the potentially available pool in academic medicine is African-American. So improving our "race numbers" in today's world means stealing somebody from another school; this "zero sum game" is not particularly productive.

It's important to figure out the particular diversity problems and opportunities that you face. Sometimes you have a pipeline issue (too few people in the applicant pool), and sometimes a different issue (such as not doing a good job of nurturing and advancing the careers of people where they are). Issues are different when selecting someone for a job or training opportunity, versus supporting someone who is already a member of your community. Not only does our school have very few African-American faculty members, but our satisfaction survey suggests that we could do a better job of making those that we do have feel more welcome and supported.

Whatever your official role, you should strive to treat each individual well. Stereotypes can sometimes help



alert you to the fact that a person might have a problem outside your experience. For example, Mary told us about a student who couldn't imagine adults who don't have a driver's license, or any other government-sponsored ID. Yet 10% or more of Americans of voting age don't have such credentials.

Don't assume that you know what people's problems are. Try to learn what they need and how you can be of help. Stereotypes cause harm when they lead you to pre-judge people as a substitute for thoughtful and supportive interactions.

**Golbeck:** Jon, you were president of the ASA. Voluntary associations are different from the organizations we are employed in. What is our professional association, the ASA, doing about implicit bias?

**Kettenring:** Well, actually, I had no clue. So I asked our esteemed Executive Director Ron Wasserstein what the ASA was doing. I'm going to share his answer with you.

The ASA policy on implicit bias goes back to November of 2011. The board of directors adopted a policy very similar to one that was already in place by the Mathematical Association of America. The title is, "Avoiding Implicit Bias: Guidelines for ASA Awards Committees" [1]. So the focus at that time was on the awards component.

I think it's encouraging that the ASA has been using this policy in a broader sense. Ron made the comment to me that we're also applying it to appointments on committees. ASA President-Elect Jessica Utts sent me a note saying she was well aware of the policy, and she is keeping these concerns in mind when making her appointments for 2016.

Here are the 12 components of the policy [1]:

1. Appoint diverse selection committees and chairs.
2. Generate a large and diverse pool of nominees.
3. Publicize the award among underrepresented groups.
4. Periodically review and discuss practices for building a pool of nominees.
5. Periodically review the description and guidelines for the award.
6. Discuss the process and criteria.
7. Consider those with strong qualifications but whose work may be less widely known.
8. Make a personal list of top nominees.  
[This avoids the undue influence of one member and insures that the list of viable nominees is as large as possible before discussion begins.]
9. Create short lists via inclusive methods.
10. Ensure every committee member's voice is heard.

11. Take adequate time to make a decision.

[This reminds me again of the book, *Thinking Fast and Slow*. Take the time to do it right.]

12. Avoid conflict of interest.

I was pleased to see that there was a very thoughtful list here. Taking a list like this and adapting it to other contexts could help give you a running start on dealing with implicit bias issues in somewhat different contexts.

## 5. What are specific success stories or best practices for achieving equal professional opportunity for all?

**Golbeck:** One example of a success is the AAUP Higher Education Salary Evaluation Kit [27], conceptualized in the early 1970s when Mary Gray, our panelist, was chair of the AAUP Committee on Women in the Profession. UC-Berkeley Professor Elizabeth L. Scott developed the kit, and Mary helped oversee this effort. The kit influenced salary adjustments.

Let's take a look at what's happening now. Arlene, what are some specific success stories or best practices for achieving equal professional opportunity for all?

**Ash:** Many have studied this key question: What works to change people's attitudes and, ultimately, their behaviors? Sadly, most of what people have tried doesn't work. But at least we're doing research on how to do better.

There are a few notable success stories. Most don't pertain to our profession. One is how symphony orchestras are no longer all male, because auditions are held with people walking on in their socks (to eliminate the sound of walking in high heels), sitting down behind a screen, and playing [14]. Those who make the judgments have to do it simply on how the music sounds, without recourse to irrelevant information like the musician's height, weight, race or sex. Unfortunately, what makes for a good faculty hire is more nuanced and multidimensional – it can't be fully blinded. We know that the hiring committee should make a list of the key job requirements and tabulate the factual evidence as to how well each candidate's credentials fit those requirements, but – unlike a symphony orchestra – we cannot use a fully blinded process to eliminate our biases.

Some success stories do apply to professions like ours. For example, five or six years ago all of the panels and all of the speakers at the microbiology professional meetings were men, even though about half the

microbiologists were women [17]. Then they changed the rules. On every committee putting together an invited session there has to be a woman; there has to be someone who pays attention to the gender makeup. In just a few years, they achieved equity.

So there are things we can do. We must go for the “low hanging fruit” and adopt process changes that have been shown to be useful. We must also continue to learn more about what strategies work, and continue to collect and report data that reveal the extent to which we are meeting our legal and ethical responsibility to treat everyone fairly.

**Golbeck:** Judy, what are specific success stories in convincing colleagues, leaders, and decision-makers that their judgments may also be affected by implicit bias?

**Singer:** There is an example about the Harvard statistics department, where I earned my PhD in 1983. The department was founded in 1958. There was a 58-year gap, but I’m very pleased to say that, as of Thursday, the Harvard statistics department now has two tenured women, biostatisticians Xihong Lin and Xiaole (Shirley) Liu who previously were appointed to the Harvard biostatistics department.

What changed? First, we have a committed dean who told the department that it could not conduct any searches unless they changed their ways. This requires honest conversations, something that can be hard to do. But if you don’t have those honest conversations, people are going to keep thinking that what they are doing is just fine. Second, this dean appointed a new department chair who is committed to addressing their issues. His name is Neil Shepard. He was recruited to Harvard just three years ago. He was appointed department chair starting in September, but he’s already done a number of things. He began by having honest conversations with his dean, with me, and most importantly with his departmental colleagues. They reviewed the lack of success in recruiting women. He was particularly concerned about their all-male search committees, which was inevitable if all the voting members of the department are male. Now that Lin and Liu have tenured appointments in the department – which means they get to vote on tenured appointments and can serve on search committees – I’m optimistic about the future.

The dean has now given the ok to do searches, and he’s agreed that the department can make not just one hire, but several. Research shows that when you have to make decisions one by one they are much more likely to favor the white or, in the case of statistics, Asian man. If there is a cluster-hire, the committee is unlikely

to come up with a pool of three white or Asian men. It wouldn’t be acceptable and they know it. We’ve also rewritten the language for the ad to say, “The department is keenly interested in diversifying its faculty”. This language is a sign that they want to do things differently. So a dean and a department chair can make a huge difference.

If your departments look anything like the Harvard statistics department for the prior 58 years (and the rosters of your departments indicate that some of yours do), then having these kinds of honest conversations and bringing the issues out into the open makes an enormous difference. Change is possible. Leadership is key.

**Golbeck:** Judy, Harvard didn’t appoint a man and a woman this year, they appointed two women. Why is it helpful to have two women brought in to the department?

**Singer:** When you’re the only woman in the room, often times you don’t get paid attention, sometimes you’re asked to bring the coffee, a number of things happen. So having two women makes an enormous difference. You’re not alone. You don’t necessarily have to see eye to eye with the other woman, but you’re no longer a token. Of course, having more is also better: two is better than one, but three is better than two.

**Golbeck:** Then what happens when women lead or take responsibility, Mary?

**Gray:** Better things happen than when they don’t. That doesn’t mean good things always happen. If you have women in leadership positions, they’re sometimes as insensitive to the issues as men, either because they just don’t see them, or because they have a so-called Queen Bee complex: “You know, I like being the only woman, I’m so much better than the rest of the people who are the worker bees, and I’d like to keep it that way”. So you can’t count on women doing better. But the importance of there being more than one minority person in a group has been shown with a lot of jury research, where when you only have one person on a jury who is holding out, it is very difficult to be heard. If you have two it’s much easier. Now obviously if you have three it’s easier still.

In fact, there is a lot of research on what constitutes a sufficient cluster to make a difference. For example, in the affirmative action lawsuit at the University of Texas at Austin, which is now at the Supreme Court once again, the issue is: How many minorities do you need to get at the university in order for them actually to have an effect? We’ll see what the court says this time, because you can’t count on the Supreme Court either.

But, that's another way women take leadership. Maybe we would be unhappy in the case of Texas because it is a woman who sued Texas, but there have been some very important reforms based on a woman suing. We've had pension reform because a woman, Natalie Norris, allowed herself to be discriminated against just long enough to file suit. Lilly Ledbetter is another example where the law got changed. These things happen.

You can see some changes. I worked on the committee that came up with the 12 statements on implicit bias for the ASA that Jon presented. We worked very hard to come up with the statements. The excuse that the ASA used for the underrepresentation of women among research award winners is one many organizations use: All of these awards committees are different committees. There are seven different research awards given by ASA. Each one has a separate committee. So it doesn't seem to them that bad that they come up with men all or almost all of the time. But when you put together the effects of all of these committees, then you see the total effect. Part of the solution was supposed to be not only that the new policy was going to be followed, but that there would be an oversight awards committee that would meet with the people who were on the other awards committees and see that we got more equity. At one JSM, I gave a presentation to the awards committee about why this was important. I wish I could say I made a great impact, and we then suddenly had lots more women getting research awards. That's probably not the case.

The one thing that I think is really effective in scientific organizations is the thing Arlene mentioned about the microbiologists. When you get a woman to organize the panel, you are much more likely to get women on the panel. Simply increasing the number of women who are doing the invited panels or sessions, or who are doing an activity where you have to choose other people, is something that has managed to work. So that's something I can recommend to everyone, whether it is a solution to the problem of implicit bias or not. And when Harvard manages to get women into tenured positions in the math department, then I'll believe they've solved the problem!

**Singer:** I didn't say I had solved the problem! We had one woman in mathematics. She was recruited to Princeton to become the first woman in their math department.

**Gray:** And then she moved again.

**Singer:** Yes, this is what happens. Faculty members get poached. The market for top talent is absolutely crazy.

Earlier we were talking about the gender imbalance in the distribution of ASA awards. One way to solve this imbalance is to ask committees to nominate several outstanding candidates, not just one. Ask each committee to nominate three potential recipients and then all the committee nominations go to a central oversight committee. All the committees are unlikely to nominate three white or Asian men, especially if you bring the gender imbalance to their attention. And if all seven awards committees do, then the problem is *much* more serious and needs a more major intervention.

**Golbeck:** Talking about serious problems, pay inequity is a persistent serious problem. Nick, you have spent a lot of time thinking about pay equity.

**Jewell:** There was a very interesting article in *The New York Times* about five years ago illustrating levels of pay inequity in various professions [8]. Some professions display serious pay inequities. The problem gets much worse the higher the median salary; if you're a woman, you're going to tend to do much worse relatively speaking as you get promoted. That's the reality.

The article shows there are two professions where women get paid more 'on average' than men: Special education teachers and postal clerks. Women do equally badly in academe and law: Women professors and women lawyers both get paid 20% below men.

Aren't universities supposed to be progressive? Aren't all university faculty members supposed to be socialists? So how can this happen? In reality, universities are one of the most conservative professions on the pay issue. Every one of you who work in a university will go to work next Monday in an institution which is exploiting women every single day at least in terms of their pay. It's a fundamental problem.

President John F. Kennedy signed the equal pay act in 1963. That was 52 years ago. The modern version of this act is the paycheck fairness act. It states that women should get equal pay for equal work. This act has been blocked in Congress in 2010, 2012, and 2014. It hasn't even come up for a vote. I hope we're successful in 2016, but since 2016 is a presidential election year, it's probably not going to happen.

So what can we do? The first thing is to be active 'politically', because then things will change. If you're not political about anything, choose something to be political about. Locally, do a pay equity study. Just do one in your department. You can usually get this information. In my institution, it's on the web. Or ask your dean or your chair to do a pay equity study. Demand it. In fact, there are laws that obligate institutions to do pay equity studies. And you're also obligated by law

to do something about it if you determine the existence of a pay gap. So the way to get something done is to actually start the process.

Because we are statisticians, we can provide tools. You've heard about Betty Scott's AAUP tools for conducting salary equity studies. Provide the tools to chairs and deans so they can identify inequities. Many people are determining salaries as part of their jobs, and they are not looking at the data systematically. They are doing the 'every single act separately' rather than looking at the groups. When you're looking at salaries as a whole in a department or a college, you see the inequities more clearly. You've got to keep challenging the status quo.

Women have to raise their voices and take action. But women shouldn't be the only ones: Men should raise their voices and take action, too. The pay situation is unacceptable. This is a moral issue that women are not being paid equally. There are other things about part time labor and flexibility in the workplace for women that are similarly very important.

**Golbeck:** Nick, do you want to say something about gated communities?

**Jewell:** The term 'gated communities' makes us think of all these country club places where people live. The term was coined by sociologist Robert Hironimus-Wendt in a broader sense. Some academic disciplines are gated communities. Let me explain.

A new study at UC-Berkeley ranked every department by the fraction of their faculty that are women [23]. Mechanical engineering has the highest fraction of male faculty, followed by electrical engineering, computer science, physics, economics, and mathematics. Statistics is ranked 14th. The women are teaching all of the wonderful arts that are so important and vibrant for any university. Art history, Spanish and Portuguese, theater/dramatic performance, art, and film and media have the highest fraction of female faculty.

There is no reason why men and women should have equal interests. But who gets paid the most? It's the mechanical engineer, not the art history professor. Why is that, and why is average faculty pay so correlated with the ranking? Well, you could argue, maybe the mechanical engineers are more important because they build bridges, and so they are more important to society; art history doesn't contribute much to the economy. Is this really the only reason behind the noticeable pay variation?

These pay equity reports come out, and people explain them away. For example, they say that we ought

to stratify on department or discipline. If you leave economics in a social sciences gender comparison and don't adjust for it, the imbalance is usually huge. Economists get paid far more than anyone else in the social sciences, and they are almost all men. So people take economics out of the social sciences comparisons, and then they say: That's better now, there is less inequity.

But is it really better? Perhaps women are being blocked from being economists, systematically. It's a gated community. Who lets you into the economics department where the salaries are big? The economics department lets you in. Who are the economics department? They're predominantly men. So some disciplines are essentially gated communities. I think it's not enough to require that in the French department a man and woman should be paid the same and stop there. I think we should have comparable pay for comparable work. Professors do the same job, pretty much, if they're in the French department or the statistics department, and we ought to get away from this idea that certain professions are very highly paid at the same time making them almost exclusively male.

## 6. How can you start a conversation about diversity in your organization?

**Golbeck:** In order to effect change, it's important not just to start, but also to sustain conversations about diversity in our organizations. Arlene mentioned that what you would do depends upon your vantage point. One of the things I was able to do when I worked at the higher education system level in Kansas was to stimulate diversity through the department chairs at the seven state universities. Kansas had an annual statewide higher education conference on diversity and multiculturalism. The state also had an annual statewide professional development conference for university department chairs. The university provosts and I got together, and we said: "This Michael Tillford Conference on Diversity and Multiculturalism, it seems like the same people go every time and we are preaching to the choir. Maybe we should devote the department chairs conference to diversity this year". And we did. We focused on the role of the department chair as a diversity agent. We didn't want the effort to end with the conference, so we asked the department chairs to formulate action plans that they would take back to their campuses. That was all done in the presence of the department chairs' deans and provosts. It was something we were able to do at the system level.

There are all kinds of things happening now on our

campuses and in our organizations. Let's see some examples of what has been done from those vantage points. Marcia, how can you start a conversation about diversity in your organization?

**Gumpertz:** I think having discussions about diversity is key to change. People need to engage with these ideas. It's not something that statisticians or people in professional organizations or university departments usually talk about, or think is part of what they should be talking about, so it might be hard to get the discussion started. But it's important, because in a university department, faculty are the ones who make the hiring and admissions decisions. People in the department or unit set the climate for everyone there. So it's really important to have a way for people to start thinking about these ideas.

I have a few thoughts about how to get discussions started. With statisticians, data are very powerful: Showing trends, demographics, results of climate studies, results of any salary equity studies, etc. These results are a great way to get the discussion started.

You may find that it's helpful to have an outside facilitator come in and lead the discussion. People aren't used to thinking diversity is something that the organization needs to talk about, or people are afraid of saying the wrong thing. There are barriers to talking about race, religion, gender, and all of these things. It's very tricky. So having a trained facilitator can be a good thing.

Another easy way to get faculty talking about diversity is to have a reading group about some interesting topic, like the use of SAT's in admissions for students. A book came out this summer, *The Tyranny of the Meritocracy*, which can generate lots of great discussion [16]. So a reading group is also a good thing.

It's important to be able to articulate the important questions for your organization. What things affect your mission that people will care about in your organization? This will help you get buy-in for having these discussions. At our physics department retreat this summer, I asked the faculty to anonymously send stories beforehand of things they had either experienced themselves, or students had experienced, or they had witnessed, or microaggressions they themselves had perpetrated. Then we had a session built around these stories. People were incredulous: This happened in our department? Knowing that the stories were from their own colleagues made it real to them and important to talk about. Everyone brainstormed in small groups about how you would handle different situations. If this came up in your class, how would you

handle this? If this is going on in the department, what can we do as a department to change and prevent this kind of thing from happening again?

I think for academic groups, having discussions around topics from faculty's own experience is a really good way to start discussions. I also think that showing real data – not just from the whole institution or country, but from that department – is very effective.

**Golbeck:** Nick, briefly, how important is it to have men engaged in these conversations in all aspects and not leave the work to diversity committees and specified diversity representatives?

**Jewell:** I think it's very important to engage men in these conversations. I went to the recent Women in Statistics Conference. There were about three men and 600 women attending. I recommend the conference if it is done again. It was a very interesting experience for a man to be such an underrepresented minority; men almost never have that kind of experience. It's extremely important to have champions of change. Men run the world, unfortunately or fortunately, depending on your perspective. Men make the rules, and if you don't get the foxes involved in 'changes', the chicken house is going to get raided forever. We have to have men involved, and we have to challenge men to be champions of change.

**Golbeck:** Judy, what are some effective strategies for enlisting *men* (i.e., not "just" women) in being champions for change?

**Singer:** I'll start out with an anecdote. A number of years ago I was on a diversity committee at the university, and it was all women. At the end of the year, the chair of the committee asked, "What can we do to improve this committee next year?" I said, "Invite some men". One of the women on the committee said, "Why would we want to do that"? I looked at her and said, "That's what they used to say about having women on committees"!

I disagree with the notion that diversity work should be done solely by women and minorities. Fortunately, there are now generational shifts. Look to some of your younger colleagues, who are more likely to be concerned about these issues. Younger men were brought up in a different generation. I've been doing events for recently tenured faculty. The men are keenly interested in gender issues. They don't want to be in departments that misbehave anymore.

Expertise also matters. Reach out to the social scientists on your campus. Some of what we've been discussing is standard social psychology and behavioral economics; reach out to your colleagues in those departments. You undoubtedly have local experts who can be effective champions on this point. Bring in out-

side people to talk and get a conversation going on your campus.

Push for gender-neutral family-friendly policies. We want all of our faculty members to be able to balance work and life, not just the women. This will make it better for everybody.

Our recent Harvard faculty climate survey included two standard items that I'd like to mention [31]. The first is: Does your school or department make genuine efforts to recruit female faculty? We found that women are three times as likely to disagree with the statement. The other item is: "I feel the climate for female faculty in your department is as good as it is for male faculty". We found women are twice as likely to disagree. When men see these results, it's hard to argue that everything is fine, and it really becomes quite compelling.

We also looked at data on time use for our tenure track (junior) faculty. Instead of just looking at gender differences, which is what typically has been done, we broke it out by family household composition: No partner or children; partner, no children; children and nonworking partner; and children and working partner or single. We found no significant difference in hours at work for men and women in any of these categories. For women and men who don't have children and don't have partners, we found indistinguishable differences in hours on household duties. We did, however, find a 20-hour gap in the amount of time spent on housework and family responsibilities for the category of people with children and working partner or single. The men are actually doing more – they're doing 20 hours a week on average – but the women are doing 40 hours a week on top of their average 60-hour work-week. In retrospect, I'm sorry I didn't ask about how many hours they sleep and what kinds of leisure activities they pursue!

These results have catalyzed conversations. They've been picked up in newspapers and also in our departments, where colleagues are saying: I didn't actually realize that our female faculty members with children effectively have another full time job at home.

**Gray:** I want to bring up one thing about how to start conversations. You probably can't see, but I'm wearing a 3/4 of a Euro pin that a group of feminists gave me when I was at a conference in Helsinki about an opera by a woman composer and woman mathematician. Three quarters of a Euro is what the average salary is for women in Europe compared with men. (And I'm having one made with the Sacagawea dollar!)

**Golbeck:** Jon, briefly, what are some of the measurement issues?

**Kettenring:** Actually, there's been quite a bit of discussion about measurement issues. It's pretty clear we

know how to count things, and we know how to take a look at the numbers and assess their strengths and weaknesses. However, the more I've thought about this question, the more it seems to me that some of the things we really want to measure are pretty difficult. I've looked at a lot of the literature. I don't find much encouraging there. But it doesn't mean we should give up. It just means we should recognize that, for example, just because we have a few women on the faculty, that's not the end of the story. There are a lot of deeper issues that we all are aware of, I think, that are hard to quantify. It reminds me of a statement that Deming used to make a lot about complicated situations [7]: The things we really want to know are unknown and unknowable. But we shouldn't give up because of that.

## 7. Conclusion

**Golbeck:** Implicit bias affects decisions to pursue career tracks [25], and it affects levels of success within careers [34]. John Lennon famously once said in an interview that every child is an artist until he's told he's not an artist. Arlene, what happens in life to some of our creative talents?

**Ash:** The bulk of Pollack's book describes the multitude of ways in which her incredible talent, drive and determination to become a physicist was ultimately overcome, and her reflection and research on why that happened, and what has and hasn't changed since she was in school. The larger point of these various threads is that our society frequently fails to nurture, and often actively discourages, the creative talents of many, many people.

**Golbeck:** Mary, you've had so many years of experience in this whole area of gender equity. Do you have any concluding remarks?

**Gray:** The conversation needs to get started, and you need to be persistent. Sometimes it makes you unpopular. I don't think we talked about that, but it can happen. People don't always like to have things pointed out for which they are responsible and probably should not have done. But you need to stick with it. You need to be persistent. And you need to involve men. One of the things I've found is that men suddenly become much more interested in these problems when their daughters are job-seeking age. But get them started a little bit earlier than that. You need a lot of allies.

**Golbeck:** We asked ourselves two basic questions when we decided to propose this JSM session. (1) Can

the statistics profession afford to have unintentional perceptions continue to hinder the success of women and other underrepresented groups in statistics? (2) Is it reasonable for our profession to continue to underuse its human capital toward the advancement of science? As our good colleague Dan Solomon points out, diversity in the workplace isn't just an issue of ethics; it is also an economic, academic, and business imperative [32].

## Acknowledgments

The authors wish to thank Fritz Scheuren for inviting us to prepare the content of our JSM Invited Session for publication. Amanda Golbeck wishes to thank the Waterman Fund of the Philadelphia Foundation for its support of the Elizabeth L. Scott Projects.

## References

- [1] American Statistical Association (ASA), Avoiding Implicit Bias: Guidelines for ASA Awards Committees, Accessed on 11/16/2015 at <http://www.amstat.org/awards/acguidelines.cfm>.
- [2] Association for Women in Science (AWIS), Implicit bias in STEM. Accessed on 11/12/2015 at: [https://awis.site-ym.com/?Awards\\_Outcomes](https://awis.site-ym.com/?Awards_Outcomes).
- [3] M.R. Banaji and A.G. Greenwald, *Blindspot: Hidden Biases of Good People*, Delacorte Press (2013). See: <https://implicit.harvard.edu/implicit/>.
- [4] M. Bertrand and S. Mullainathan, Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination, *American Economic Review* **94**(4) (2004), 991–1013.
- [5] M. Carnes, S. Geller, E. Fine, J. Sheridan and J. Handelsman, NIH Director's pioneer awards: Could the selection process be biased against women? *Journal of Womens Health* **14**(8) (2005), 684–691.
- [6] J. Carpenter, Google's algorithm shows prestigious job ads to men, but not to women, Here's why that should worry you, *The Washington Post*, 6 July 2015.
- [7] W.E. Deming, *Out of the Crisis*, The MIT Press, 2000.
- [8] H. Fairfield and G. Roberts, Why is her paycheck smaller? *The New York Times*, 18 May 2010. Accessed on 10/11/15 at: [http://www.nytimes.com/interactive/2009/03/01/business/20090301\\_WageGap.html?\\_r=0](http://www.nytimes.com/interactive/2009/03/01/business/20090301_WageGap.html?_r=0).
- [9] M. Gladwell, *Blink: The Power of Thinking Without Thinking*, Back Bay Books, 2007.
- [10] A.L. Golbeck, Where are the women in the JSM Registration Guide? *AmStat News* **421** (2012), 16–17.
- [11] A.L. Golbeck, *Equivalence: Elizabeth L. Scott at Berkeley*, Chapman & Hall/CRC Press, forthcoming.
- [12] A.L. Golbeck and C.A. Molgaard, Professional awards in statistics: Chipping away at gender disparities in the USA, *Proceedings of the 59<sup>th</sup> International Statistical Institute World Statistics Congress*, Hong Kong, China, 2013, pp. 5155–5160.
- [13] A.L. Golbeck, I. Olkin and Y.R. Gel, eds, *Leadership and Women in Statistics*, Chapman and Hall/CRC Statistics Press, 2016.
- [14] C. Goldin and C. Rouse, Orchestrating impartiality: The impact of "blind" auditions on female musicians, *American Economic Review* **90**(4) (2000), 715–741.
- [15] M. Gray, Leadership and the legal system, in: *Leadership and Women in Statistics*, A.L. Golbeck, I. Olkin and Y.R. Gel, eds, Chapman and Hall/CRC Press, 2016, pp. 229–244.
- [16] L. Guinier, *The Tyranny of the Meritocracy: Democratizing Higher Education in America*, Beacon Press, 2015.
- [17] S. Jaschik, Ending all-male panels, *Inside Higher Ed* (5 August 2015), Accessed on 11/16/2015 at <https://www.insidehighered.com/news/2015/08/05/microbiologists-show-its-possible-achieve-gender-equity-scholarly-presentations#.VclcOGe3Nqs.mailto>.
- [18] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2013.
- [19] C.C. Miller, Algorithms and bias: Q and A with Cynthia Dwork, *The New York Times*, 10 August 2015. Accessed on 9/2/2015 at <http://mobile.nytimes.com/2015/08/11/upshot/algorithms-and-bias-q-and-a-with-cynthia-dwork.html>.
- [20] L. Mlodinow, *Subliminal: How Your Unconscious Mind Rules Your Behavior*, Vintage Books, 2012.
- [21] C.A. Moss-Racusin, J.F. Dovidio, V.L. Brescoll, M.J. Graham and J. Handelsman, Science faculty's subtle gender biases favor male students, *Proceedings of the National Academy of Sciences* **109**(41) (2012), 16474–16479.
- [22] Office of the Senior Vice Provost, *Best Practices for Conducting Faculty Searches*, Harvard University, 2016. Available for download at [www.faculty.harvard.edu/FacultySearch](http://www.faculty.harvard.edu/FacultySearch).
- [23] Office of the Vice Provost for the Faculty, Report on the UC Berkeley Faculty Salary Equity Study, January, 2015, p. 16. Accessed on 10/11/2015 at: <http://vpf.berkeley.edu/sites/default/files/Equity%20Study%20Report%20final%201-26-15%20-revised.pdf>.
- [24] Ohio State University, Kirwan Institute for the Study of Race and Ethnicity, 2015. Accessed on 11/12/2015 at <http://kirwaninstitute.osu.edu/>.
- [25] E. Pollack, *The Only Woman in the Room: Why Science is Still a Boy's Club*, Beacon Press, 2015.
- [26] D.L. Sackett, Bias in analytic research, *Journal of Chronic Diseases* **32** (1979), 51–63.
- [27] E.L. Scott, *Higher Education Salary Evaluation Kit*, AAUP, 1977.
- [28] E.L. Scott, Letter to NSF fellowship program officers, *Elizabeth L. Scott papers*, Bancroft Library, UC-Berkeley, 1984.
- [29] K. Servick, Forensic labs explore blind testing to prevent errors, *Science* **31** (July 2015), 462–463.
- [30] J.M. Sheltzer and J.C. Smith, Elite male faculty in the life sciences employ fewer women, *Proceedings of the National Academy of Sciences* **111**(28) (2014), 10107–10112.
- [31] J.D. Singer, Fostering the advancement of women in academic statistics, in: *Leadership and Women in Statistics*, A.L. Golbeck, I. Olkin and Y.R. Gel, eds, Chapman and Hall/CRC Press, 2016, pp. 413–428.
- [32] D.L. Solomon, If you would consider a woman..., in: *Leadership and Women in Statistics*, A.L. Golbeck, I. Olkin and Y.R. Gel, eds, Chapman and Hall/CRC Press, 2016, pp. 395–411.
- [33] L.M. Wah, *The Color of Fear* (video), Stirfry, 1994.
- [34] J. Wolfers, Even famous female economists get no respect, *The New York Times* (11 November 2015). Accessed on 11/16/2015 at <http://www.nytimes.com/2015/11/12/upshot/even-famous-female-economists-get-no-respect.html>.

# Power and Privilege: Reshaping the Opportunity Structure for Equitable Leadership in Statistics and Data Science



Emma Benn

If my cup won't hold but a pint, and yours holds a quart,  
wouldn't you be mean not to let me have my little half measure  
full?—Sojourner Truth, *Ain't I a Woman*, 1851 (McKissack,  
1992)

**Abstract** Innovation in statistics and data science requires the equitable elevation of diverse perspectives. In this chapter, I explore the mechanisms by which the tightly woven interplay between power and privilege serves as an impediment to and catalyst for progressively reshaping the opportunity structure for leadership in our field.

## Introduction

When Dr. Amanda Golbeck first approached me to contribute my thoughts around power and privilege to this book, I was quite hesitant for a number of reasons. First, I had never written a book chapter and while I have a wife, a novelist, who can churn out 20 chapters per day, or at least it seems that way, the proposition still felt a bit daunting to me. Second, while most of my service activities have been devoted to increasing diversity and inclusion in the field of (bio)statistics and data science, I questioned whether I had been in the game long enough to have sufficient insight to share with others. I could think of many prominent statisticians who would be better suited for this important discourse, but Dr. Golbeck, who I had first met when we participated on an *Implicit Bias in the Profession of Statistics* panel at the 2017 Joint Statistical Meetings (JSM), thought otherwise. She effectively persuaded me to put my self-doubt aside, rise to the challenge, and lend my perspective to this necessary conversation.

---

E. Benn (✉)

Icahn School of Medicine at Mt. Sinai, Brooklyn, NY, USA

e-mail: [emma.benn@mountsinai.org](mailto:emma.benn@mountsinai.org)

© Springer Nature Switzerland AG 2021

A. L. Golbeck (ed.), *Leadership in Statistics and Data Science*,  
[https://doi.org/10.1007/978-3-030-60060-0\\_2](https://doi.org/10.1007/978-3-030-60060-0_2)

19



## Defining Power and Privilege: Starting with #MeToo

So what would be the best way to initiate a conversation about power and privilege in statistics? Perhaps the starting point lies within the larger society and the current cultural shifts we have all been observing in real-time as a direct result of the “me too” Movement™. This grassroots movement, also referred to all over social media as the #MeToo movement, has been particularly instrumental in empowering survivors of sexual violence as well as directly confronting and holding accountable the sociostructural and sociopolitical systems at the global level that directly and/or indirectly promote sexual violence (Me too Movement, <https://metoomvmt.org/about/#history>). In a recent TEDWomen 2018 address (Me Too is a movement, not a moment, [https://www.ted.com/talks/tarana\\_burke\\_me\\_too\\_is\\_a\\_movement\\_not\\_a\\_moment?language=en](https://www.ted.com/talks/tarana_burke_me_too_is_a_movement_not_a_moment?language=en)), Ms. Tarana J. Burke, civil rights activist and founder of the “me too” Movement™, made a particularly informative statement about the relationship between power and privilege in our society by suggesting that “...anybody in a position of power comes with privilege and it renders those without power more vulnerable.” She further elaborated that “power and privilege doesn’t always have to destroy and take, it can be used to serve and build.” Ms. Burke’s eloquently articulated interlinking of power, privilege, and vulnerability served as a much-needed catalyst for me in that she provided a contemporary conceptual framework from which to draw for this important discourse.

Yet, before I can delve deeper, I find it necessary to make sure that I am working with standard definitions of each of these constructs so that it is easier to discuss their interplay. According to Merriam-Webster (power, <https://www.merriam-webster.com>; privilege, <https://www.merriam-webster.com>), *power* refers to having the “ability to act or produce an effect” or to having “possession of control, authority, or influence over others,” whereas *privilege* is defined as “a right or immunity granted as a peculiar benefit, advantage, or favor.” To be *vulnerable* (vulnerable, <https://www.merriam-webster.com>) would imply that an individual is “capable of being physically or emotionally abused” or, more generally, “open to attack or damage.” Thus, my interpretation of Ms. Burke’s first statement is that there are endowed benefits for individuals who are granted control or authority over others and this can leave those under their control at risk for harm. For example, I can remember being told as a graduate student by a professor that I had been awarded a fellowship that was linked to both my educational endeavors and my underrepresented status as a racial/ethnic minority. I did not know I had been nominated for anything, so this of course was an exciting surprise. Yet, the professor went further to request that I give my fellowship away to support another student in my program who while we were similarly pursuing a graduate degree, the student was not underrepresented. I was devastated. This professor, who I looked up to and who I thought had my best interest in mind, attempted to exploit my underrepresented status to benefit someone else. It was, as my godmother Nancy Kirby would say, “a soul wound.” Eventually, when I finally gathered enough courage, I spoke to someone in the administration who I thought might be able

to advocate for me and fortunately the situation was resolved. Yet, I must say that for quite a time thereafter I feared that there might be repercussions for speaking up. While that was the first time in my career that I worried about advocating for myself, which entailed pushing back against what I felt was a misuse of power or inappropriate treatment by a superior, it certainly was not the last time.

## Shining a Light on the “Vulnerable”

The harm or “damage” that vulnerable members of our community may experience as an intentional or unintentional consequence of the power and privilege exerted by another statistician or data scientist may come in many forms. In my case, something that I rightfully deserved could have been taken away from me. Speaking out, however, is not always a feasible option for some statisticians and data scientists. For example, I have known some individuals who, due to the fact that their visa status is directly linked to their employer, cannot speak out when they are mistreated. They also cannot easily walk away from their job when they feel exploited by their supervisor or overworked. Rarely do our discussions around diversity and inclusion tap into the challenges that these immigrants confront in our field. Perhaps this stems from the fact that we typically associate vulnerability with having less representation. However, nonresident aliens—a term the National Center for Education Statistics Integrated Postsecondary Education Data System (Definitions for new race and ethnicity categories, <https://nces.ed.gov/ipeds/report-your-data/race-ethnicity-definitions>) uses to refer to “a person who is not a citizen or national of the United States and who is in this country on a visa or temporary basis and does not have the right to remain indefinitely”—represent a similar if not greater proportion of those getting graduate degrees in our field, at least since 2011, as compared to US citizens and permanent residents (Statistics and biostatistics degree data, <https://www.amstat.org/asa/education/Statistics-and-Biostatistics-Degree-Data.aspx?hkey=0a32a96f-2f47-4d67-b91e-0b329f93eece>; Highlights from 2017 Degree release: Bachelor’s numbers close in on Master’s, <https://magazine.amstat.org/blog/2018/08/01/2017-degree-report/>). Therefore, we must make additional efforts to ensure that the needs of this vulnerable subgroup of immigrants are not overlooked.

Yet, before I move on, I think it is necessary to pause and reflect on this term, nonresident alien, which definitely does not sit well with me and I am certain does not sit well with the immigrants who are assigned this label with such negative connotation. Our language used to categorize different subgroups in our statistics and data science community, unfortunately, has the potential to strip individuals of their dignity. While referring to a specific group of immigrants in our field as nonresident aliens may be one major example of this dilemma, another example could be the damage that may be caused, even if unintentional, when we address someone by something other than their preferred pronouns or when we assume heteronormativity. Thornton, Green, and Benn (2019) posit that statisticians and

data scientists must acknowledge their latent assumptions about gender conformity and heteronormativity if we are to proactively create a more inclusive culture. Some ASA conferences have also started to provide individuals with the opportunity to place their preferred pronouns on their registration badges, but even outside of those spaces, those statisticians and data scientists in positions of power and privilege must find a way to use language in a way that is inviting and inclusive, rather than further marginalizing already vulnerable groups.

Students are also particularly vulnerable due to the fact that they are highly dependent on their professors and mentors to successfully meet the expectations of their educational programs as well as to jumpstart their professional careers. Thus, when a professor or mentor “harms” a student, this can result in serious repercussions for the student that may also negatively impact them as professionals. Harm, in this case, can come in many forms. Recently, statisticians, women in particular, have used their personal blogs (Lum, 2017; McClure, 2017) to go public about past experiences of sexual harassment and assault as students by professors or other prominent statisticians within the university setting as well as at statistics conferences. I should not have to explain why sexual harassment and assault is a clear abuse of power and privilege, especially when the perpetrator is not held accountable, and why this could be traumatic and likely damaging over the long term for a student, or anyone regardless of their educational or professional status. By serving with a diverse group of statisticians on the ASA Task Force on Sexual Harassment and Assault (<https://ww2.amstat.org/committees/commdetails.cfm?txtComm=ABTBOD05>; Final report of the Task Force 22jan2019, <https://www.amstat.org/asa/files/pdfs/Task%20Force%20on%20Sexual%20Harassment%20and%20Assault-FinalReport.pdf>), under the leadership of Dr. Leslie McClure, we proactively worked over the course of about 1 year to: (1) assess the magnitude of sexual harassment and assault in the ASA community, (2) revise the ASA Activities Conduct Policy, and (3) make evidence-based, policy-level recommendations for the confidential reporting of incidences of sexual harassment and assault as well as for a comprehensive review and unbiased adjudication of complaints of sexual misconduct. The work of the Task Force was instrumental in moving the statistical community forward, opening up important discourse around sexual misconduct in our field, as well as identifying proactive, progressive steps our community can take with regard to creating a safer, more inclusive, and equitable culture, especially for the most vulnerable, including students, women regardless of professional rank, and other groups.

However, there are other types of “harm” students may experience that, at times, may go unnoticed. This can take the form of exploitation when students are not credited for their work or pushed beyond their limits. In 2018, the Guardian published a piece in their Academics Anonymous series entitled *Academia is built on exploitation. We must break this vicious circle* (<https://www.theguardian.com/higher-education-network/2018/may/18/academia-exploitation-university-mental-health-professors-plagiarism>) in which the anonymous author posited that even well-intentioned professors can exploit their graduate students due to the overwhelming expectations that can come with the job. The author explained:

Even a well-meaning academic has to juggle many responsibilities. For example, a research professor must teach hundreds of undergraduates, supervise a team of PhD researchers, manage research grants and collaborations and edit publications and dissertations. Very few professors are able to accomplish all this without the overwhelming burden affecting their character and judgement.

Unfortunately, the ones most likely to be on the receiving end of this are the PhD students, postdoctoral researchers and non-academic staff, who are often on short employment contracts. These people are in very tenuous positions and have little if any recourse.

The author further scrutinized the academic system that at times both normalizes and ignores the negative behaviors of professors as it relates to the intentional or unintentional exploitation of students. While this is a global problem and thus not localized to statistics, it still speaks to the power and privilege that professors have over students and the fact that this problem is systematically reinforced by unrealistic institutional and/or self-imposed expectations around productivity and promotion. The unfair and inequitably distributed expectations are exacerbated for faculty of color due to their “token” status and female faculty for whom the number of service-related obligations and required time commitment to these tasks often surpass that which is expected of our White and/or male colleagues. Yet, we are still expected to meet or exceed the productivity of our less tasked counterparts (Edwards & Ross, 2018; Hanasono et al., 2019; Settles, Buchanan, & Dotson, 2019).

In other circumstances, professors or mentors could unfairly halt a student’s progress through their program or to the next phase as a professional due to reasons that are completely unrelated to the student’s performance or demonstrated understanding of statistics or data science, but rather due to a personal conflict or the student’s refusal to acquiesce to some exorbitant request by a professor or mentor. I can think of several statisticians that have had a professor, mentor, or supervisor unfairly attempt to prevent their progress forward as a student or professional, and while most would not openly share their experiences in public, their silence should not be interpreted as the “harm” being forgotten.

## **Power and Privilege as a Catalyst for Transformation**

Yet, while it is important to call attention to any misuse of power and privilege in our field and in society more broadly, Ms. Burke also provides us with an optimistic outlook on the transformative capacity that those with power and privilege possess. Thus, statisticians with power can use their privilege to strategically restructure our field in a more altruistic and equitable manner such that particular subgroups with less power and privilege, and most likely more vulnerable, are more optimally positioned for success and leadership.

Restructuring our field to promote equitable success and leadership can take very different forms. Sometimes it can happen informally. For example, when I think back over my approximately 7-year career as a professor, the majority of the invitations I have gotten to give a lecture, be on a panel, serve on a committee, etc.

at least in the first 5 years or so of my career have most often been due to women at more senior ranks, mostly women of color and/or LGBTQ, inviting me. I keep wondering if it was that others in the field did not see me or notice my potential, or perhaps it stems from the fact that individuals from underrepresented groups are keenly aware that we must strategically and deliberately work to position each other so that eventually others will see us. It was because of these individuals that my visibility increased. For example, it was Dr. Kim Sellers who invited me to be on the aforementioned *Implicit Bias* panel at JSM, as a follow-up to a panel she had invited me to serve on at the Women in Statistics and Data Science (WSDS) conference, and it is meeting Dr. Golbeck at that JSM panel and subsequent work I have done around increasing diversity which positioned me to lend my voice as a contributor to this book. Yet, it should not solely be the responsibility of underrepresented groups in our field to increase our own visibility and eventually gain access to opportunities that we might not have gotten otherwise.

## Intersectional Progress

I think that many of our current leaders in statistics realize that they have the power and privilege to create a more progressive opportunity structure, rather than follow the status quo. In other words, some of our leaders have begun to look to see who is not represented that should be represented before they recommend someone for a leadership position, or to give a talk, serve on a committee, serve as a journal editor, etc. There is still a lot more work to be done, however. This stems from the fact that initiatives to create a more equitable opportunity structure for leadership and success can unintentionally benefit the most privileged of the disadvantaged. I came across this phenomenon when I was introduced to the work of Kimberle Crenshaw, a professor of law and civil rights advocate, in preparation for a 2017 WSDS panel entitled *Is There Room for Intersectional Feminism in Statistics and Data Science* with Dr. Mine Dogucu and Dr. Maria Garcia. Crenshaw (1989) provides a very illustrative analogy in arguing the need for intersectionality, or intersectional feminism more specifically, when she states the following:

Imagine a basement which contains all people who are disadvantaged on the basis of race, sex, class, sexual preference, age and/or physical ability. These people are stacked . . . — with those on the bottom being disadvantaged by the full array of factors, up to the very top, where . . . all those disadvantaged by a singular factor brush up against the ceiling. A hatch is developed through which those placed immediately below can crawl . . . [it] is generally available only to those who—due to the singularity of their burden . . . are in the position to crawl through. Those who are multiply-burdened are generally left below . . .

Thus, our leaders' efforts to reshape the opportunity structure in our field will be most successful if they are intersectional. For example, efforts in our field to ensure gender equity in leadership must make sure that they not only benefit white women, but also benefit lesbian or gender nonconforming women or women with disabilities or women of color or immigrant women or others who as Crenshaw



explains are not disadvantaged by a singular factor, but rather by multiple factors. As an African-American, lesbian, woman, academic, and biostatistician, this idea of intersectionality is particularly important to me. I think, however, that as statisticians and data scientists who have intersectional identities are increasingly represented in leadership, structural-level initiatives aimed at creating a more equitable opportunity structure will undoubtedly benefit a broader, more diverse group of individuals.

## **The Power and Privilege Continuum**

However, we should not assume that only those with the most power and privilege are responsible for reshaping the opportunity structure in statistics and data science. It is important to acknowledge that power and privilege exist on a continuum rather than in some absolute fashion. In other words, each of us needs to identify our own power and privilege if we are to proactively shift the opportunity structure for leadership in our field. For example, as a first-year doctoral student, I found the fact that I was the only student of color in my graduate biostatistics program to be quite unsettling. I felt that it was not about our inability to do biostatistics, but more about the lack of exposure of underrepresented groups to the field itself. While I did not believe that I could wave some magic wand to fix this problem, I did think that collectively with other students and with the support of the administration, faculty across multiple programs, and departmental staff, we might be able to change the demographic landscape of the graduate program.

That is exactly why I cofounded the Biostatistics Epidemiology (formerly Enrichment) Summer Training (BEST) Diversity Program in 2008 at Columbia University's Mailman School of Public Health. The 8-week summer program exposes diverse (i.e., racial/ethnic minorities, economically disadvantaged, or those with disabilities) undergraduates to biostatistics and its application to cardiovascular disease-related research as well as prepare students for graduate school success. I, along with other graduate students, taught in the program, mentored the undergraduates, and assisted with many of the administrative logistics. After the first year of the program, Dr. Melissa Begg, who had been very supportive of our program and who had taken on a leadership role, was able to secure funding for BEST from the National Heart Lung and Blood Institute of the National Institutes of Health. Initially, BEST was not an easy feat. In fact, looking back, I do not know if I would advise students to take on such a big undertaking as it was difficult at times balancing the rigorous expectations of my doctoral program with ensuring that BEST ran smoothly. Yet, I was extremely passionate about BEST and when you believe in something so much, I guess you find a way to make everything work. The program has been extremely successful and has lasted for over a decade. By 2017, BEST had trained approximately 110 students of which 75% of those who had graduated from college had gone on to pursue graduate degrees (Simply the best: Celebrating a decade of diversity, <https://www.mailman.columbia.edu/public-health-now/news/simply-best-celebrating-decade-diversity>).

I really could not foresee the major impact that this program would have on diversifying the field at the time that we created it; however, I think it is a good example of understanding that our voice and our efforts can be influential regardless of where we sit on the power and privilege ladder. Even if we are not fully equipped to change the opportunity structure on our own, working collectively with others who have varying levels of power and privilege has the potential to have a more expansive impact than waiting for those at the top to figure out what strategy to pursue. I am very appreciative of the support and encouragement that Dr. Begg gave me at such an early stage in my biostatistics career as well as her continued encouragement of my pursuits around diversity and inclusion to this day. I try to pay that encouragement forward, especially when I see students and early-career professionals, like Jemar Bather and Dr. Suzanne Thornton, who have passionately worked to change the demographic landscape of our field for racial/ethnic minorities and LGBTQ+ statisticians and data scientists, respectively, and who have, quite honestly, challenged midlevel and senior-level professionals to reconsider whether we are actually doing enough.

## **Final Thoughts**

In summary, I would like to offer some final thoughts that are not aimed at concluding this discussion, but are instead aimed at propelling this conversation forward. If we are truly serious about ensuring that diverse statisticians and data scientists are more equitably promoted to leadership positions within our field, then we will have to directly confront the manner by which we intentionally or unintentionally assign power and privilege to some and vulnerability to others. The burden associated with making this necessary cultural shift should not rest solely on the most vulnerable. Instead, those of us who are most comfortably situated along the power and privilege continuum should be simultaneously contributing to and leading efforts that: (1) hold those accountable who misuse their power and privilege, (2) empower the least privileged among us, and (3) proactively advocate for a progressive, intersectional, reconfiguration of the opportunity structure in our field, so that we can collectively make progress together. These efforts should not go unnoticed. Instead, they should be celebrated as it is the dynamic cultural shifts resulting from the elevation of diverse perspectives that will undoubtedly lead to methodologic innovation, rather than the replication of the contributions of a privileged minority.

## Action Points

- Reflect on each of our unique positions within the power and privilege continuum and the extent to which we can individually and collectively reshape the opportunity structure for visibility, advancement, and leadership within our field.
- Develop and implement structural-level policies and initiatives in statistics and data science that promote the empowerment of all members of our community and can counteract the systematic intentional and unintentional mechanisms by which exploitation becomes inevitable.
- Ensure that opportunities for empowerment and advancement in statistics and data science are intersectional in nature, and thus will more optimally benefit vulnerable subgroups with varying levels of privilege.

## References

- Academics Anonymous: Academia is built on exploitation. We must break this vicious circle* [Internet]. 2018 [cited 29 August 2019]. Retrieved from <https://www.theguardian.com/higher-education-network/2018/may/18/academia-exploitation-university-mental-health-professors-plagiarism>.
- ASA Task Force on Sexual Harassment and Assault [Internet]. [cited 29 August 2019]. Retrieved from <https://ww2.amstat.org/committees/commdetails.cfm?txtComm=ABTBOD05>.
- Crenshaw, K. (1989). *Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics* (pp. 139–167). University of Chicago Legal Forum.
- Definitions for new race and ethnicity categories* [Internet]. [cited 26 August 2019]. Retrieved from <https://nces.ed.gov/ipeds/report-your-data/race-ethnicity-definitions>.
- Edwards, W. J., & Ross, H. H. (2018). What are they saying? Black faculty at predominantly white institutions of higher education. *Journal of Human Behavior in the Social Environment*, 28(2), 142–161.
- Final report of the Task Force 22jan2019* [Internet]. 2019 [cited 29 August 2019]. Retrieved from <https://www.amstat.org/asa/files/pdfs/Task%20Force%20on%20Sexual%20Harassment%20and%20Assault-FinalReport.pdf>.
- Hanasono, L. K., Broido, E. M., Yacobucci, M. M., Root, K. V., Peña, S., & O'Neil, D. A. (2019). Secret service: Revealing gender biases in the visibility and value of service. *Journal of Diversity in Higher Education*, 12(1), 85–98.
- Highlights from 2017 Degree release: Bachelor's numbers close in on Master's* [Internet]. 2018 [cited 26 August 2019]. Retrieved from <https://magazine.amstat.org/blog/2018/08/01/2017-degree-report/>.
- Lum K. *Medium* [Internet] 2017. [cited 29 August 2019]. Retrieved from <https://medium.com/@kristianlum/statistics-we-have-a-problem-304638dc5de5>.
- McClure L. *Stat Girl Blog* [Internet] 2017. [cited 2019]. Retrieved from <https://statgirlblog.wordpress.com/2017/04/07/sexual-assault-and-harassment-in-stem-we-can-no-longer-afford-to-be-silent/>.
- McKissack, P. (1992). *Sojourner truth : Ain't I a woman?* New York Scholastic Paperbacks.
- Me Too is a movement, not a moment* [Video file]. Retrieved from [https://www.ted.com/talks/tarana\\_burke\\_me\\_too\\_is\\_a\\_movement\\_not\\_a\\_moment?language=en](https://www.ted.com/talks/tarana_burke_me_too_is_a_movement_not_a_moment?language=en) [Internet]. [cited 9 September 2019].
- Me too Movement. Retrieved from <https://metoomvmt.org/about/#history>.



- power*. [Internet]. 2019 [cited 26 August 2019]. Retrieved from <https://www.merriam-webster.com>.
- privilege*. [Internet]. 2019 [cited 26 August 2019]. Retrieved from <https://www.merriam-webster.com>.
- Settles, I. H., Buchanan, N. T., & Dotson, K. (2019). Scrutinized but not recognized: (In)visibility and hypervisibility experiences of faculty of color. *Journal of Vocational Behavior*, 113, 62–74.
- Simply the best: Celebrating a decade of diversity* [Internet]. 2017 [cited 29 August 2019]. Retrieved from <https://www.mailman.columbia.edu/public-health-now/news/simply-best-celebrating-decade-diversity>.
- Statistics and biostatistics degree data* [Internet]. 2019 [cited 26 August 2019]. Retrieved from <https://www.amstat.org/asa/education/Statistics-and-Biostatistics-Degree-Data.aspx?hkey=0a32a96f-2f47-4d67-b91e-0b329f93eece>.
- Thornton, S., Green, B., & Benn, E. (2019). Friends and allies: LGBT+ inclusion in statistics and data science. *Significance*, 16(3), 39–41.
- vulnerable*. [Internet]. 2019 [cited 26 August 2019]. Retrieved from <https://www.merriam-webster.com>.

## THE JEDI CORNER

# At the Rise of JEDI: Lessons Learned from Fall of the Jedi Order in *Star Wars*

*The JEDI Corner is a regular component of Amstat News in which statisticians write about and educate our community about JEDI-related matters. If you have an idea or article for the column, email JEDI Outreach Group member Cathy Furlong at [communicate@datascijedi.org](mailto:communicate@datascijedi.org).*



**Shu-Min Liao** is an assistant professor of statistics at Amherst College. She completed her PhD training in statistics at The Pennsylvania State University in 2009, but her professional career was unfortunately interrupted by life-threatening medical complications in 2013–2018. She is currently living with multiple disabilities and minority identities—all of which make her a better educator and researcher. Sparked by those unusual experiences, Shu-Min is passionate about STEM Education research and antiracism work (besides discrete copula modeling). She is a member of the ASA JEDI Outreach Group and a faculty facilitator for the “Being Human in STEM” course at Amherst College.

This summer, I enjoyed the fun of watching all three *Star Wars* movies in chronological order with my nine-year-old boy. He is finally old enough to enjoy those films with me; plus, not being able to travel during the pandemic gave us time (or excuses?) to watch all the movies on nine Friday nights.

Far from being a *Star Wars* fan, I was delighted to see the new ASA outreach group named JEDI—an abbreviation for Justice, Equity, Diversity, and Inclusion. But when I told my son (shortly after JSM 2021) that his mom wants to be a JEDI, he responded without even blinking his eyes, “Nah, I want to be Darth Vader, or Kylo Ren, but not a Jedi; the Jedi failed anyway.” [Disclosure: We hadn’t watched the last film, *The Rise of Skywalker*, at that point.]

Was my boy right? If so, why did the Jedi (Order) fail?

A few quick Google searches led me to hundreds, if not thousands, of discussions about this topic. While, unsurprisingly, different people had different views and interpretations of why the Jedi failed, what astonished me the most was the similarities between the problems the Jedi were facing in the *Skywalker* saga and those we are facing in academia today.

Let me explain.

**Hierarchical Power Structure and Privilege:** As the CBR.com article, “The 15 Most Evil Things About the Jedi Order,” stated, “The pyramidal hierarchy of the Jedi Order is topped with the extremely powerful Jedi High Council. ... With long-lived members like Yoda and Plo Koon, the Order could potentially be ruled by the same individuals for centuries, greatly limiting the *diversity* and *natural evolution* of Jedi leadership with the times.” Only the most powerful Jedi could join the Council, which had the power and privilege to decide the destinies of other Jedi (including who/when to be *promoted* as a Jedi Master) and the fate of the galaxy.

Such hierarchy is nothing new in academia. As Brian Martin pointed out in Chapter 3 of *Tied Knowledge: Power in Higher Education*, “Universities are hierarchical. ... The academic hierarchy, like other hierarchies, is a system in which people exercise power not by virtue of their personal talents but by virtue of the position they occupy.”

More recently, Emma Benn in *Power and Privilege: Reshaping the Opportunity Structure for Equitable Leadership in Statistics and Data Science* provided an incisive description of the relationship between power and privilege: “There are endowed benefits for individuals who are granted control or authority over others and this can leave those under their control at risk for harm.”

Both statements sparked me to ask: Who has the power and privilege in our system? Who is at risk? How can we use power and privilege to “serve and build,” rather than “destroy and take”?

**Problematic Dichotomy:** The Jedi are light and good, while the Sith are dark and bad. Nothing in between. But aren’t light and dark “two sides of the same coin”? No dark, no light. The Jedi arrogantly claimed they were the sole gatekeeper of the Force, but who decided the Jedi’s perspectives about the Force were more “correct” than those of the Sith’s?

In academia, similar dichotomy is everywhere. For example, here is some advice I’ve gotten from senior colleagues over the years:

“You should try to work on more popular topics; otherwise, it would be very hard for you to get grants.”

“You should focus on methodological research, rather than pedagogical research, as statistics education research would only count as real research if it’s funded.”

“Antiracism and JEDI works are important, but unfortunately you won’t get tenured with them.”

Getting more publications and grants is good; others are subordinate. Research is more important than teaching and service. People with tenure are more successful (and have more voice and power) than those without. How many of us ever questioned whose criteria is good, who defines success, who decides which types of scholarship or research are more significant than others?

#### **Rigid System with Little Room for Humanity or Failure:**

The Jedi Order throughout the prequel-era was dogmatic and used rigid criteria to decide who was Force-sensitive and who wasn't. They believed 8- or 9-year-old kids were too old to begin Jedi training. Most Jedi were recruited when they were toddlers or little kids and forbidden to have emotional bonds (like love) or negative feelings (like anger), as those human emotions were viewed as the passageway to the Dark side. How dehumanizing!

If you believe our academic system is better than the Jedi's, please think again. For most tenure-track positions, you only have five years to establish a new, independent, and successful research agenda, so make sure you choose your research topics carefully, especially if you don't have a 'powerful' academic family to pave the road for you. Moreover, make no mistake, as you probably won't have enough time to start a new research line from scratch if you fail. Failure is unfortunately forbidden. Long vacations are missions impossible. Remember to bring your research with you on all family trips. Don't think about self-care, as that is just a waste of time. Oh, also reveal no weakness or tears in front of students and colleagues, as those are signs of losers. [Okay, I'm exaggerating here, but you get the point.]

How about our students? The structural inequity and injustice they face is no less—it might be more—than junior faculty, especially for those from equity-seeking groups. According to Joe Feldman, author of *Grading for Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms*, students are on the bottom of the academic power hierarchy (especially those coming from disadvantaged families) and have little say in what should be taught in class and how, even though it's for their learning.

Most colleges and universities, at best, provide equal, but not necessarily equitable, resources to students, which penalizes those arriving to college with inequitable preparation along with a grading system designed with the privileged in mind.

Dichotomy is everywhere in student life, too—good vs. bad grades, winners vs. losers, popular vs.

### **Further Reading**

Groneman, Thomas. 2017. Blight Side of the Force: The 15 Most Evil Things About the Jedi Order. *CBR.com*.

<https://bit.ly/3pCGmvQ>

Martin, Brian. 1998. Hierarchy. In *Tied Knowledge: Power in Higher Education*.

<https://bit.ly/3rRYdBy>

Benn, Emma. 2021. *Power and Privilege: Reshaping the Opportunity Structure for Equitable Leadership in Statistics and Data Science*.

<https://bit.ly/3EWh0zH>

Propelled by Possibility: Tarana Burke Speaks at TEDWomen 2018

<https://bit.ly/31Q3crt>

Feldman, Joe. 2019. *Grading for Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms*.

<https://bit.ly/3oGx5Ue>

Murakami, Haruki. 2009. Always on the Side of the Egg. *Haaretz.com*.

<https://bit.ly/3rSu0lR>

nerd, rich vs. poor, white vs. non-white, etc. Further, higher education rarely prepares students for unavoidable failures and challenges in life. We do little to help young adults understand their unique identities and strengths or to learn self-care skills for promoting life-long mental health. Some might argue it's not the faculty's job, but shouldn't we—the educators—educate not only students' brains, but also their hearts?

Solutions? My favorite novelist, Haruki Murakami, said a decade ago:

**We are all human beings**, individuals transcending nationality and race and religion, and we are all fragile eggs faced with a solid wall called The System. To all appearances, we have no hope of winning. The wall is too high, too strong, and too cold. If we have any hope of victory at all, it will have to come from our believing in the *utter uniqueness and irreplaceability* of our own and others' souls and from our believing in the warmth we gain by *joining souls together*. ... The System did not make us; we made the System.

May the Force be with us! ■

# Reading Unit 2:

## Study Design

(Estimated Reading Time: 2 hours,  
Complete Reading by 8am 01/31/2022))

**UO21:** Through readings, group discussion and reflection, students will explain through specific real-life examples how individual racism and discrimination can lead scientists to make unethical decisions around data that contribute to structural oppression of marginalized groups. (CO 11)

**UO22:** Through readings, group discussion and reflection, students will analyze how power that scientists hold in society interacts with societal stereotypes about marginalized groups. (CO 11)

1. Brandt, A. M. (1978). Racism and research: the case of the Tuskegee Syphilis Study. *Hastings center report*, 21-29.
2. Henry, B. V., Chen, H., Edwards, M. A., Faber, L., & Freischlag, J. A. (2021). A New Look at an Old Problem: Improving Diversity, Equity, and Inclusion in Scientific Research. *The American Surgeon*, 00031348211029853.
3. Lewis, J. E., DeGusta, D., Meyer, M. R., Monge, J. M., Mann, A. E., & Holloway, R. L. (2011). The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias. *PLoS Biology*, 9(6), e1001071.
4. Antrosio, J. (2011). "Mismeasuring Gould in 'The Mismeasure of Science.'" *Living Anthropologically* website, <https://www.livinganthropologically.com/mismeasuring-gould/>. First posted 14 June 2011. Revised 7 September 2017.
5. Anderson, M., & Fienberg, S. E. (1999). Who counts?: The politics of census-taking in contemporary America. *Russell Sage Foundation*, 11-34.
6. Anderson, M. (2020). The Census and the Japanese" Internment": Apology and Policy in Statistical Practice. *Social Research: An International Quarterly*, 87(4), 789-812.

# Reading Unit 2: Discussion Questions

1. What are some examples of structural violence against African-Americans and Japanese-Americans from these readings? How are these examples related to the field of statistics?
2. The fifth reading, *from the book Who Counts? The politics of census-taking in contemporary American*, introduces the idea of the differential undercount. What is a differential undercount, and how is it related to the DEI concept of structural oppression?

# Reading Unit 2: Reflection Questions

1. In this reflection we will be exploring our social identities as we consider how society is structured. Complete the following chart about your own identity. You do not have to share this identity in your written reflection. After mapping your social identity, think about the readings and the experiences of the people that were or are discriminated against. How are their social identities similar or different than yours?
2. How did the readings make you feel about the structural inequities in society that result from flawed data collection processes?
3. What responsibility do you feel toward making sure scientific and governmental data collection efforts are equitable?

<b>Domain</b>	<b>My Identity</b>	<b>At Slippery Rock University, this identity gives me a position of: Privilege (P), Marginalization (M)</b>	<b>In the place I grew up, this identity gave me a position of: Privilege (P), Marginalization (M)</b>
<b>Race</b> (e.g., white, black, biracial)			
<b>Ethnicity</b> (e.g., German, Polish, Cuban, Hispanic, Japanese)			
<b>Gender identity/ expression</b> (e.g., woman, man, nonbinary, transgender male, transgender female)			
<b>Sexual orientation</b> (e.g., homosexual, heterosexual, bisexual, pansexual)			
<b>Religion</b> (e.g., Catholic, Protestant, Jewish, Muslim)			
<b>Dis/ability</b> (e.g., able-bodied, autistic, ADHD, deaf)			
<b>First Language</b> (e.g., English, Spanish)			
<b>First Generation Status</b> (first generation in college in nuclear family, not first generation in college)			
<b>Socio-Economic Status</b> (as you feel it is best defined – use of governmental assistance programs such as WIC, Medicaid, or SNAP; middle class, working class, etc.)			
<b>Other:</b> (you define)			

# Racism and Research: The Case of the Tuskegee Syphilis Study

by ALLAN M. BRANDT

In 1932 the U.S. Public Health Service (USPHS) initiated an experiment in Macon County, Alabama, to determine the natural course of untreated, latent syphilis in black males. The test comprised 400 syphilitic men, as well as 200 uninfected men who served as controls. The first published report of the study appeared in 1936 with subsequent papers issued every four to six years, through the 1960s. When penicillin became widely available by the early 1950s as the preferred treatment for syphilis, the men did not receive therapy. In fact on several occasions, the USPHS actually sought to prevent treatment. Moreover, a committee at the federally operated Center for Disease Control decided in 1969 that the study should be continued. Only in 1972, when accounts of the study first appeared in the national press, did the Department of Health, Education and Welfare halt the experiment. At that time seventy-four of the test subjects were still alive; at least twenty-eight, but perhaps more than 100, had died directly from advanced syphilitic lesions.<sup>1</sup> In August 1972, HEW appointed an investigatory panel which issued a report the following year. The panel found the study to have been "ethically unjustified," and argued that penicillin should have been provided to the men.<sup>2</sup>

This article attempts to place the Tuskegee Study in a historical context and to assess its ethical implications. Despite the media attention which the study received, the HEW *Final Report*, and the criticism expressed by several professional organizations, the experiment has been largely misunderstood. The most basic questions of *how* the study was undertaken in the first place and *why* it continued for forty years were never addressed by the HEW investigation. Moreover, the panel misconstrued the nature of the experiment, failing to consult important documents available at the National Archives which bear significantly on its ethical assessment. Only by examining the specific ways in which values are engaged in scientific research can the study be understood.

## Racism and Medical Opinion

A brief review of the prevailing scientific thought regarding race and heredity in the early twentieth century is fundamental for an understanding of the Tuskegee Study. By the turn of the century, Darwinism had provided a new ration-

ale for American racism.<sup>3</sup> Essentially primitive peoples, it was argued, could not be assimilated into a complex, white civilization. Scientists speculated that in the struggle for survival the Negro in America was doomed. Particularly prone to disease, vice, and crime, black Americans could not be helped by education or philanthropy. Social Darwinists analyzed census data to predict the virtual extinction of the Negro in the twentieth century, for they believed the Negro race in America was in the throes of a degenerative evolutionary process.<sup>4</sup>

The medical profession supported these findings of late nineteenth- and early twentieth-century anthropologists, ethnologists, and biologists. Physicians studying the effects of emancipation on health concluded almost universally that freedom had caused the mental, moral, and physical deterioration of the black population.<sup>5</sup> They substantiated this argument by citing examples in the comparative anatomy of the black and white races. As Dr. W. T. English wrote: "A careful inspection reveals the body of the negro a mass of minor defects and imperfections from the crown of the head to the soles of the feet. . . ."<sup>6</sup> Cranial structures, wide nasal apertures, receding chins, projecting jaws, all typed the Negro as the lowest species in the Darwinian hierarchy.<sup>7</sup>

Interest in racial differences centered on the sexual nature of blacks. The Negro, doctors explained, possessed an excessive sexual desire, which threatened the very foundations of white society. As one physician noted in the *Journal of the American Medical Association*, "The negro springs from a southern race, and as such his sexual appetite is strong; all of his environments stimulate this appetite, and as a general rule his emotional type of religion certainly does not decrease it."<sup>8</sup> Doctors reported a complete lack of morality on the part of blacks:

Virtue in the negro race is like angels' visits—few and far between. In a practice of sixteen years I have never examined a virgin negro over fourteen years of age.<sup>9</sup>

A particularly ominous feature of this overzealous sexuality, doctors argued, was the black males' desire for white women. "A perversion from which most races are exempt," wrote Dr. English, "prompts the negro's inclination towards white women, whereas other races incline towards females of their own."<sup>10</sup> Though English estimated the "gray matter of the negro brain" to be at least a thousand years behind that of the white races, his genital organs were overdeveloped. As Dr. William Lee Howard noted:

The attacks on defenseless white women are evidences of racial instincts that are about as amenable to ethical culture as is the inherent odor of the race. . . . When education will

ALLAN M. BRANDT is a doctoral candidate in the Department of History, Columbia University. He is presently writing a social history of venereal disease in the United States. Mr. Brandt was a student intern at The Hastings Center in 1977.

reduce the size of the negro's penis as well as bring about the sensitiveness of the terminal fibers which exist in the Caucasian, then will it also be able to prevent the African's birthright to sexual madness and excess.<sup>11</sup>

One southern medical journal proposed "Castration Instead of Lynching," as retribution for black sexual crimes. "An impressive trial by a ghost-like kuklux klan [sic] and a 'ghost' physician or surgeon to perform the operation would make it an event the 'patient' would never forget," noted the editorial.<sup>12</sup>

According to these physicians, lust and immorality, unstable families, and reversion to barbaric tendencies made blacks especially prone to venereal diseases. One doctor estimated that over 50 percent of all Negroes over the age of twenty-five were syphilitic.<sup>13</sup> Virtually free of disease as slaves, they were now overwhelmed by it, according to informed medical opinion. Moreover, doctors believed that treatment for venereal disease among blacks was impossible, particularly because in its latent stage the symptoms of syphilis become quiescent. As Dr. Thomas W. Murrell wrote:

They come for treatment at the beginning and at the end. When there are visible manifestations or when harried by pain, they readily come, for as a race they are not averse to physic; but tell them not, though they look well and feel well, that they are still diseased. Here ignorance rates science a fool. . . .<sup>14</sup>

Even the best educated black, according to Murrell, could not be convinced to seek treatment for syphilis.<sup>15</sup> Venereal disease, according to some doctors, threatened the future of the race. The medical profession attributed the low birth rate among blacks to the high prevalence of venereal disease which caused stillbirths and miscarriages. Moreover, the high rates of syphilis were thought to lead to increased insanity and crime. One doctor writing at the turn of the century estimated that the number of insane Negroes had increased thirteen-fold since the end of the Civil War.<sup>16</sup> Dr. Murrell's conclusion echoed the most informed anthropological and ethnological data:

So the scourge sweeps among them. Those that are treated are only half cured, and the effort to assimilate a complex civilization driving their diseased minds until the results are criminal records. Perhaps here, in conjunction with tuberculosis, will be the end of the negro problem. Disease will accomplish what man cannot do.<sup>17</sup>

This particular configuration of ideas formed the core of medical opinion concerning blacks, sex, and disease in the early twentieth century. Doctors generally discounted socioeconomic explanations of the state of black health, arguing that better medical care could not alter the evolutionary scheme.<sup>18</sup> These assumptions provide the backdrop for examining the Tuskegee Syphilis Study.

### The Origins of the Experiment

In 1929, under a grant from the Julius Rosenwald Fund, the USPHS conducted studies in the rural South to determine the prevalence of syphilis among blacks and explore the

possibilities for mass treatment. The USPHS found Macon County, Alabama, in which the town of Tuskegee is located, to have the highest syphilis rate of the six counties surveyed. The Rosenwald Study concluded that mass treatment could be successfully implemented among rural blacks.<sup>19</sup> Although it is doubtful that the necessary funds would have been allocated even in the best economic conditions, after the economy collapsed in 1929, the findings were ignored. It is, however, ironic that the Tuskegee Study came to be based on findings of the Rosenwald Study that demonstrated the possibilities of mass treatment.

Three years later, in 1932, Dr. Taliaferro Clark, Chief of the USPHS Venereal Disease Division and author of the Rosenwald Study report, decided that conditions in Macon County merited renewed attention. Clark believed the high prevalence of syphilis offered an "unusual opportunity" for observation. From its inception, the USPHS regarded the Tuskegee Study as a classic "study in nature,"\* rather than an experiment.<sup>20</sup> As long as syphilis was so prevalent in Macon and most of the blacks went untreated throughout life, it seemed only natural to Clark that it would be valuable to observe the consequences. He described it as a "ready-made situation."<sup>21</sup> Surgeon General H. S. Cumming wrote to R. R. Moton, Director of the Tuskegee Institute:

The recent syphilis control demonstration carried out in Macon County, with the financial assistance of the Julius Rosenwald Fund, revealed the presence of an unusually high rate in this county and, what is more remarkable, the fact that 99 per cent of this group was entirely without previous treatment. This combination, together with the expected cooperation of your hospital, offers an unparalleled opportunity for carrying on this piece of scientific research which probably cannot be duplicated anywhere else in the world.<sup>22</sup>

Although no formal protocol appears to have been written, several letters of Clark and Cumming suggest what the USPHS hoped to find. Clark indicated that it would be important to see how disease affected the daily lives of the men:

The results of these studies of case records suggest the desirability of making a further study of the effect of untreated syphilis on the human economy among people now living and engaged in their daily pursuits.<sup>23</sup>

It also seems that the USPHS believed the experiment might demonstrate that antisyphilitic treatment was unnecessary. As Cumming noted: "It is expected the results of this study may have a marked bearing on the treatment, or conversely the non-necessity of treatment, of cases of latent syphilis."<sup>24</sup>

The immediate source of Cumming's hypothesis appears to have been the famous Oslo Study of untreated syphilis. Between 1890 and 1910, Professor C. Boeck, the chief of the

---

\*In 1865, Claude Bernard, the famous French physiologist, outlined the distinction between a "study in nature" and experimentation. A study in nature required simple observation, an essentially passive act, while experimentation demanded intervention which altered the original condition. The Tuskegee Study was thus clearly not a study in nature. The very act of diagnosis altered the original conditions. "It is on this very possibility of acting or not acting on a body," wrote Bernard, "that the distinction will exclusively rest between sciences called sciences of observation and sciences called experimental."



Oslo Venereal Clinic, withheld treatment from almost two thousand patients infected with syphilis. He was convinced that therapies then available, primarily mercurial ointment, were of no value. When arsenic therapy became widely available by 1910, after Paul Ehrlich's historic discovery of "606," the study was abandoned. E. Bruusgaard, Boeck's successor, conducted a follow-up study of 473 of the untreated patients from 1925 to 1927. He found that 27.9 percent of these patients had undergone a "spontaneous cure," and now manifested no symptoms of the disease. Moreover, he estimated that as many as 70 percent of all syphilitics went through life without inconvenience from the disease.<sup>25</sup> His study, however, clearly acknowledged the dangers of untreated syphilis for the remaining 30 percent.

Thus every major textbook of syphilis at the time of the Tuskegee Study's inception strongly advocated treating syphilis even in its latent stages, which follow the initial inflammatory reaction. In discussing the Oslo Study, Dr. J. E. Moore, one of the nation's leading venereologists wrote, "This summary of Bruusgaard's study is by no means intended to suggest that syphilis be allowed to pass untreated."<sup>26</sup> If a complete cure could not be effected, at least the most devastating effects of the disease could be avoided. Although the standard therapies of the time, arsenical compounds and bismuth injection, involved certain dangers because of their toxicity, the alternatives were much worse. As the Oslo Study had shown, untreated syphilis could lead to cardiovascular disease, insanity, and premature death.<sup>27</sup> Moore wrote in his 1933 textbook:

Though it imposes a slight though measurable risk of its own, treatment markedly diminishes the risk from syphilis. In latent syphilis, as I shall show, the probability of progression, relapse, or death is reduced from a probable 25-30 percent without treatment to about 5 percent with it; and the gravity of the relapse if it occurs, is markedly diminished.<sup>28</sup>

"Another compelling reason for treatment," noted Moore, "exists in the fact that every patient with latent syphilis may be, and perhaps is, infectious for others."<sup>29</sup> In 1932, the year in which the Tuskegee Study began, the USPHS sponsored and published a paper by Moore and six other syphilis experts that strongly argued for treating latent syphilis.<sup>30</sup>

The Oslo Study, therefore, could not have provided justification for the USPHS to undertake a study that did not entail treatment. Rather, the suppositions that conditions in Tuskegee existed "naturally" and that the men would not be treated anyway provided the experiment's rationale. In turn, these two assumptions rested on the prevailing medical attitudes concerning blacks, sex, and disease. For example, Clark explained the prevalence of venereal disease in Macon County by emphasizing promiscuity among blacks:

This state of affairs is due to the paucity of doctors, rather low intelligence of the Negro population in this section, depressed economic conditions, and the very common promiscuous sex relations of this population group which not only contribute to the spread of syphilis but also contribute to the prevailing indifference with regard to treatment.<sup>31</sup>

In fact, Moore, who had written so persuasively in favor

of treating latent syphilis, suggested that existing knowledge did not apply to Negroes. Although he had called the Oslo Study "a never-to-be-repeated human experiment,"<sup>32</sup> he served as an expert consultant to the Tuskegee Study:

I think that such a study as you have contemplated would be of immense value. It will be necessary of course in the consideration of the results to evaluate the special factors introduced by a selection of the material from negro males. Syphilis in the negro is in many respects almost a different disease from syphilis in the white.<sup>33</sup>

Dr. O. C. Wenger, chief of the federally operated venereal disease clinic at Hot Springs, Arkansas, praised Moore's judgment, adding, "This study will emphasize those differences."<sup>34</sup> On another occasion he advised Clark, "We must remember we are dealing with a group of people who are illiterate, have no conception of time, and whose personal history is always indefinite."<sup>35</sup>

The doctors who devised and directed the Tuskegee Study accepted the mainstream assumptions regarding blacks and venereal disease. The premise that blacks, promiscuous and lustful, would not seek or continue treatment, shaped the study. A test of untreated syphilis seemed "natural" because the USPHS presumed the men would never be treated; the Tuskegee Study made that a self-fulfilling prophecy.

### Selecting the Subjects

Clark sent Dr. Raymond Vonderlehr to Tuskegee in September 1932 to assemble a sample of men with latent syphilis for the experiment. The basic design of the study called for the selection of syphilitic black males between the ages of twenty-five and sixty, a thorough physical examination including x-rays, and finally, a spinal tap to determine the incidence of neuro-syphilis.<sup>36</sup> They had no intention of providing any treatment for the infected men.<sup>37</sup> The USPHS originally scheduled the whole experiment to last six months; it seemed to be both a simple and inexpensive project.

The task of collecting the sample, however, proved to be more difficult than the USPHS had supposed. Vonderlehr canvassed the largely illiterate, poverty-stricken population of sharecroppers and tenant farmers in search of test subjects. If his circulars requested only men over twenty-five to attend his clinics, none would appear, suspecting he was conducting draft physicals. Therefore, he was forced to test large numbers of women and men who did not fit the experiment's specifications. This involved considerable expense since the USPHS had promised the Macon County Board of Health that it would treat those who were infected, but not included in the study.<sup>38</sup> Clark wrote to Vonderlehr about the situation: "It never once occurred to me that we would be called upon to treat a large part of the county as return for the privilege of making this study. . . . I am anxious to keep the expenditures for treatment down to the lowest possible point because it is the one item of expenditure in connection with the study most difficult to defend despite our knowledge of the need therefor."<sup>39</sup> Vonderlehr responded: "If we could find from 100 to 200 cases . . . we would not have to do another Wassermann on useless individuals . . ."<sup>40</sup>

Significantly, the attempt to develop the sample contradicted the prediction the USPHS had made initially regarding the prevalence of the disease in Macon County. Overall rates of syphilis fell well below expectations; as opposed to the USPHS projection of 35 percent, 20 percent of those tested were actually diseased.<sup>41</sup> Moreover, those who had sought and received previous treatment far exceeded the expectations of the USPHS. Clark noted in a letter to Vonderlehr:

I find your report of March 6th quite interesting but regret the necessity for Wassermann [sic] . . . such a large number of individuals in order to uncover this relatively limited number of untreated cases.<sup>42</sup>

Further difficulties arose in enlisting the subjects to participate in the experiment, to be “Wassermanned,” and to return for a subsequent series of examinations. Vonderlehr found that only the offer of treatment elicited the cooperation of the men. They were told they were ill and were promised free care. Offered therapy, they became willing subjects.<sup>43</sup> The USPHS did not tell the men that they were participants in an experiment; on the contrary, the subjects believed they were being treated for “bad blood”—the rural South’s colloquialism for syphilis. They thought they were participating in a public health demonstration similar to the one that had been conducted by the Julius Rosenwald Fund in Tuskegee several years earlier. In the end, the men were so eager for medical care that the number of defaulters in the experiment proved to be insignificant.<sup>44</sup>

To preserve the subjects’ interest, Vonderlehr gave most of the men mercurial ointment, a noneffective drug, while some of the younger men apparently received inadequate dosages of neoarsphenamine.<sup>45</sup> This required Vonderlehr to write frequently to Clark requesting supplies. He feared the experiment would fail if the men were not offered treatment.

It is desirable and essential if the study is to be a success to maintain the interest of each of the cases examined by me through to the time when the spinal puncture can be completed. Expenditure of several hundred dollars for drugs for these men would be well worth while if their interest and cooperation would be maintained in so doing. . . . It is my desire to keep the main purpose of the work from the negroes in the county and continue their interest in treatment. That is what the vast majority wants and the examination seems relatively unimportant to them in comparison. It would probably cause the entire experiment to collapse if the clinics were stopped before the work is completed.<sup>46</sup>

On another occasion he explained:

Dozens of patients have been sent away without treatment during the past two weeks and it would have been impossible to continue without the free distribution of drugs because of the unfavorable impression made on the negro.<sup>47</sup>

The readiness of the test subjects to participate of course contradicted the notion that blacks would not seek or continue therapy.

The final procedure of the experiment was to be a spinal tap to test for evidence of neuro-syphilis. The USPHS presented this purely diagnostic exam, which often entails considerable pain and complications, to the men as a “special

treatment.” Clark explained to Moore:

We have not yet commenced the spinal punctures. This operation will be deferred to the last in order not to unduly disturb our field work by any adverse reports by the patients subjected to spinal puncture because of some disagreeable sensations following this procedure. These negroes are very ignorant and easily influenced by things that would be of minor significance in a more intelligent group.<sup>48</sup>

The letter to the subjects announcing the spinal tap read:

Some time ago you were given a thorough examination and since that time we hope you have gotten a great deal of treatment for bad blood. You will now be given your last chance to get a second examination. This examination is a very special one and after it is finished you will be given a special treatment if it is believed you are in a condition to stand it. . . .

REMEMBER THIS IS YOUR LAST CHANCE FOR SPECIAL FREE TREATMENT. BE SURE TO MEET THE NURSE.<sup>49</sup>

The HEW investigation did not uncover this crucial fact: the men participated in the study under the guise of treatment.

Despite the fact that their assumption regarding prevalence and black attitudes toward treatment had proved wrong, the USPHS decided in the summer of 1933 to continue the study. Once again, it seemed only “natural” to pursue the research since the sample already existed, and with a depressed economy, the cost of treatment appeared prohibitive—although there is no indication it was ever considered. Vonderlehr first suggested extending the study in letters to Clark and Wenger:

At the end of this project we shall have a considerable number of cases presenting various complications of syphilis, who have received only mercury and may still be considered untreated in the modern sense of therapy. Should these cases be followed over a period of from five to ten years many interesting facts could be learned regarding the course and complications of untreated syphilis.<sup>50</sup>

“As I see it,” responded Wenger, “we have no further interest in these patients *until they die*.”<sup>51</sup> Apparently, the physicians engaged in the experiment believed that only autopsies could scientifically confirm the findings of the study. Surgeon General Cumming explained this in a letter to R. R. Moton, requesting the continued cooperation of the Tuskegee Institute Hospital:

This study which was predominantly clinical in character points to the frequent occurrence of severe complications involving the various vital organs of the body and indicates that syphilis as a disease does a great deal of damage. Since clinical observations are not considered final in the medical world, it is our desire to continue observation on the cases selected for the recent study and if possible to bring a percentage of these cases to autopsy so that pathological confirmation may be made of the disease processes.<sup>52</sup>

Bringing the men to autopsy required the USPHS to devise a further series of deceptions and inducements. Wenger warned Vonderlehr that the men must not realize that they would be autopsied:

There is one danger in the latter plan and that is if the colored population become aware that accepting free hospital care means a post-mortem, every darkey will leave Macon County and it will hurt [Dr. Eugene] Dibble's hospital.<sup>53</sup>

"Naturally," responded Vonderlehr, "it is not my intention to let it be generally known that the main object of the present activities is the bringing of the men to necropsy."<sup>54</sup> The subjects' trust in the USPHS made the plan viable. The USPHS gave Dr. Dibble, the Director of the Tuskegee Institute Hospital, an interim appointment to the Public Health Service. As Wenger noted:

One thing is certain. The only way we are going to get post-mortems is to have the demise take place in Dibble's hospital and when these colored folks are told that Doctor Dibble is now a Government doctor too they will have more confidence.<sup>55\*</sup>

After the USPHS approved the continuation of the experiment in 1933, Vonderlehr decided that it would be necessary to select a group of healthy, uninfected men to serve as controls. Vonderlehr, who had succeeded Clark as Chief of the Venereal Disease Division, sent Dr. J. R. Heller to Tuskegee to gather the control group. Heller distributed drugs (noneffective) to these men, which suggests that they also believed they were undergoing treatment.<sup>56</sup> Control subjects who became syphilitic were simply transferred to the test group—a strikingly inept violation of standard research procedure.<sup>57</sup>

The USPHS offered several inducements to maintain contact and to procure the continued cooperation of the men. Eunice Rivers, a black nurse, was hired to follow their health and to secure approval for autopsies. She gave the men non-effective medicines—"spring tonic" and aspirin—as well as transportation and hot meals on the days of their examinations.<sup>58</sup> More important, Nurse Rivers provided continuity to the project over the entire forty-year period. By supplying "medicinals," the USPHS was able to continue to deceive the participants, who believed that they were receiving therapy from the government doctors. Deceit was integral to the study. When the test subjects complained about spinal taps one doctor wrote:

\*The degree of black cooperation in conducting the study remains unclear and would be impossible to properly assess in an article of this length. It seems certain that some members of the Tuskegee Institute staff such as R. R. Moton and Eugene Dibble understood the nature of the experiment and gave their support to it. There is, however, evidence that some blacks who assisted the USPHS physicians were not aware of the deceptive nature of the experiment. Dr. Joshua Williams, an intern at the John A. Andrew Memorial Hospital (Tuskegee Institute) in 1932, assisted Vonderlehr in taking blood samples of the test subjects. In 1973 he told the HEW panel: "I know we thought it was merely a service group organized to help the people in the area. We didn't know it was a research project at all at the time." (See, "Transcript of Proceedings," Tuskegee Syphilis Study Ad Hoc Advisory Panel, February 23, 1973, Unpublished typescript. National Library of Medicine, Bethesda, Maryland.) It is also apparent that Eunice Rivers, the black nurse who had primary responsibility for maintaining contact with the men over the forty years, did not fully understand the dangers of the experiment. In any event, black involvement in the study in no way mitigates the racial assumptions of the experiment, but rather, demonstrates their power.

They simply do not like spinal punctures. A few of those who were tapped are enthusiastic over the results but to most, the suggestion causes violent shaking of the head; others claim they were robbed of their procreative powers (regardless of the fact that I claim it stimulates them).<sup>59</sup>

Letters to the subjects announcing an impending USPHS visit to Tuskegee explained: "[The doctor] wants to make a special examination to find out how you have been feeling and whether the treatment has improved your health."<sup>60</sup> In fact, after the first six months of the study, the USPHS had furnished no treatment whatsoever.

Finally, because it proved difficult to persuade the men to come to the hospital when they became severely ill, the USPHS promised to cover their burial expenses. The Milbank Memorial Fund provided approximately \$50 per man for this purpose beginning in 1935. This was a particularly strong inducement as funeral rites constituted an important component of the cultural life of rural blacks.<sup>61</sup> One report of the study concluded, "Without this suasion it would, we believe, have been impossible to secure the cooperation of the group and their families."<sup>62</sup>

Reports of the study's findings, which appeared regularly in the medical press beginning in 1936, consistently cited the ravages of untreated syphilis. The first paper, read at the 1936 American Medical Association annual meeting, found "that syphilis in this period [latency] tends to greatly increase the frequency of manifestations of cardiovascular disease."<sup>63</sup> Only 16 percent of the subjects gave no sign of morbidity as opposed to 61 percent of the controls. Ten years later, a report noted coldly, "The fact that nearly twice as large a proportion of the syphilitic individuals as of the control group has died is a very striking one." Life expectancy, concluded the doctors, is reduced by about 20 percent.<sup>64</sup>

A 1955 article found that slightly more than 30 percent of the test group autopsied had died *directly* from advanced syphilitic lesions of either the cardiovascular or the central nervous system.<sup>65</sup> Another published account stated, "Review of those still living reveals that an appreciable number have late complications of syphilis which probably will result, for some at least, in contributing materially to the ultimate cause of death."<sup>66</sup> In 1950, Dr. Wenger had concluded, "We now know, where we could only surmise before, that we have contributed to their ailments and shortened their lives."<sup>67</sup> As black physician Vernal Cave, a member of the HEW panel, later wrote, "They proved a point, then proved a point, then proved a point."<sup>68</sup>

During the forty years of the experiment the USPHS had sought on several occasions to ensure that the subjects did not receive treatment from other sources. To this end, Vonderlehr met with groups of local black doctors in 1934, to ask their cooperation in not treating the men. Lists of subjects were distributed to Macon County physicians along with letters requesting them to refer these men back to the USPHS if they sought care.<sup>69</sup> The USPHS warned the Alabama Health Department not to treat the test subjects when they took a mobile VD unit into Tuskegee in the early 1940s.<sup>70</sup> In 1941, the Army drafted several subjects and told them to begin antisyphilitic treatment immediately. The

USPHS supplied the draft board with a list of 256 names they desired to have excluded from treatment, and the board complied.<sup>71</sup>

In spite of these efforts, by the early 1950s many of the men had secured some treatment on their own. By 1952, almost 30 percent of the test subjects had received some penicillin, although only 7.5 percent had received what could be considered 'adequate doses.'<sup>72</sup> Vonderlehr wrote to one of the participating physicians, "I hope that the availability of antibiotics has not interfered too much with this project."<sup>73</sup> A report published in 1955 considered whether the treatment that some of the men had obtained had "defeated" the study. The article attempted to explain the relatively low exposure to penicillin in an age of antibiotics, suggesting as a reason: "the stoicism of these men as a group; they still regard hospitals and medicines with suspicion and prefer an occasional dose of time-honored herbs or tonics to modern drugs."<sup>74</sup> The authors failed to note that the men believed they already were under the care of the government doctors and thus saw no need to seek treatment elsewhere. Any treatment which the men might have received, concluded the report, had been insufficient to compromise the experiment.

When the USPHS evaluated the status of the study in the 1960s they continued to rationalize the racial aspects of the experiment. For example, the minutes of a 1965 meeting at the Center for Disease Control recorded:

Racial issue was mentioned briefly. Will not affect the study. Any questions can be handled by saying these people were at the point that therapy would no longer help them. They are getting better medical care than they would under any other circumstances.<sup>75</sup>

A group of physicians met again at the CDC in 1969 to de-

cide whether or not to terminate the study. Although one doctor argued that the study should be stopped and the men treated, the consensus was to continue. Dr. J. Lawton Smith remarked, "You will never have another study like this; take advantage of it."<sup>76</sup> A memo prepared by Dr. James B. Lucas, Assistant Chief of the Venereal Disease Branch, stated: "Nothing learned will prevent, find, or cure a single case of infectious syphilis or bring us closer to our basic mission of controlling venereal disease in the United States."<sup>77</sup> He concluded, however, that the study should be continued "along its present lines." When the first accounts of the experiment appeared in the national press in July 1972, data were still being collected and autopsies performed.<sup>78</sup>

### The HEW Final Report

HEW finally formed the Tuskegee Syphilis Study Ad Hoc Advisory Panel on August 28, 1972, in response to criticism that the press descriptions of the experiment had triggered. The panel, composed of nine members, five of them black, concentrated on two issues. First, was the study justified in 1932 and had the men given their informed consent? Second, should penicillin have been provided when it became available in the early 1950s? The panel was also charged with determining if the study should be terminated and assessing current policies regarding experimentation with human subjects.<sup>79</sup> The group issued their report in June 1973.

By focusing on the issues of penicillin therapy and informed consent, the *Final Report* and the investigation betrayed a basic misunderstanding of the experiment's purposes and design. The HEW report implied that the failure to provide penicillin constituted the study's major ethical misjudgment; implicit was the assumption that no adequate therapy existed prior to penicillin. Nonetheless medical authorities

### Claude Bernard on Human Experimentation (1865)

Experiments, then, may be performed on man, but within what limits? It is our duty and our right to perform an experiment on man whenever it can save his life, cure him or gain him some personal benefit. The principle of medical and surgical morality, therefore, consists in never performing on man an experiment which might be harmful to him to any extent, even though the result might be highly advantageous to science, i.e., to the health of others. But performing experiments and operations exclusively from the point of view of the patient's own advantage does not prevent their turning out profitably to science. . . . For we must not deceive ourselves, morals do not forbid making experiments on one's neighbor or on one's self. Christian morals forbid only one thing, doing ill to one's neighbor. So, among the experiments that may be tried on man, those that can only harm are forbidden, those that are innocent are permissible, and those that may do good are obligatory. *Claude Bernard, An Introduction to the Study of Experimental Medicine (1865). Trans. by Henry C. Green (New York: Dover Publications, 1957).*

### From the HEW Final Report (1973)

1. In retrospect, the Public Health Service Study of Untreated Syphilis in the Male Negro in Macon County, Alabama, was ethically unjustified in 1932. This judgement made in 1973 about the conduct of the study in 1932 is made with the advantage of hindsight acutely sharpened over some forty years, concerning an activity in a different age with different social standards. Nevertheless, one fundamental ethical rule is that a person should not be subjected to avoidable risk of death or physical harm unless he freely and intelligently consents. There is no evidence that such consent was obtained from the participants in this study.

2. Because of the paucity of information available today on the manner in which the study was conceived, designed and sustained, a scientific justification for a short term demonstration study cannot be ruled out. However, the conduct of the longitudinal study as initially reported in 1936 and through the years is judged to be scientifically unsound and its results are disproportionately meager compared with known risks to human subjects involved. . . .

firmly believed in the efficacy of arsenotherapy for treating syphilis at the time of the experiment's inception in 1932. The panel further failed to recognize that the entire study had been predicated on nontreatment. Provision of effective medication would have violated the rationale of the experiment—to study the natural course of the disease until death. On several occasions, in fact, the USPHS had prevented the men from receiving proper treatment. Indeed, there is no evidence that the USPHS ever considered providing penicillin.

The other focus of the *Final Report*—informed consent—also served to obscure the historical facts of the experiment. In light of the deceptions and exploitations which the experiment perpetrated, it is an understatement to declare, as the *Report* did, that the experiment was “ethically unjustified,” because it failed to obtain informed consent from the subjects. The *Final Report's* statement, “Submitting voluntarily is not informed consent,” indicated that the panel believed that the men had volunteered for the experiment.<sup>80</sup> The records in the National Archives make clear that the men did not submit voluntarily to an experiment; they were told and they believed that they were getting free treatment from expert government doctors for a serious disease. The failure of the HEW *Final Report* to expose this critical fact—that the USPHS lied to the subjects—calls into question the thoroughness and credibility of their investigation.

Failure to place the study in a historical context also made it impossible for the investigation to deal with the essentially racist nature of the experiment. The panel treated the study as an aberration, well-intentioned but misguided.<sup>81</sup> Moreover, concern that the *Final Report* might be viewed as a critique of human experimentation in general seems to have severely limited the scope of the inquiry. The *Final Report* is quick to remind the reader on two occasions: “The position of the Panel must not be construed to be a general repudiation of scientific research with human subjects.”<sup>82</sup> The *Report* assures us that a better designed experiment could have been justified:

It is possible that a scientific study in 1932 of untreated syphilis, properly conceived with a clear protocol and conducted with suitable subjects who fully understood the implications of their involvement, might have been justified in the pre-penicillin era. This is especially true when one considers the uncertain nature of the results of treatment of late latent syphilis and the highly toxic nature of therapeutic agents then available.<sup>83</sup>

This statement is questionable in view of the proven dangers of untreated syphilis known in 1932.

Since the publication of the HEW *Final Report*, a defense of the Tuskegee Study has emerged. These arguments, most clearly articulated by Dr. R. H. Kampmeier in the *Southern Medical Journal*, center on the limited knowledge of effective therapy for latent syphilis when the experiment began. Kampmeier argues that by 1950, penicillin would have been of no value for these men.<sup>84</sup> Others have suggested that the men were fortunate to have been spared the highly toxic treatments of the earlier period.<sup>85</sup> Moreover, even these contemporary defenses assume that the men never would have been treated anyway. As Dr. Charles Barnett of Stan-

ford University wrote in 1974, “The lack of treatment was not contrived by the USPHS but was an established fact of which they proposed to take advantage.”<sup>86</sup> Several doctors who participated in the study continued to justify the experiment. Dr. J. R. Heller, who on one occasion had referred to the test subjects as the “Ethiopian population,” told reporters in 1972:

I don't see why they should be shocked or horrified. There was no racial side to this. It just happened to be in a black community. I feel this was a perfectly straightforward study, perfectly ethical, with controls. Part of our mission as physicians is to find out what happens to individuals with disease and without disease.<sup>87</sup>

These apologies, as well as the HEW *Final Report*, ignore many of the essential ethical issues which the study poses. The Tuskegee Study reveals the persistence of beliefs within the medical profession about the nature of blacks, sex, and disease—beliefs that had tragic repercussions long after their alleged “scientific” bases were known to be incorrect. Most strikingly, the entire health of a community was jeopardized by leaving a communicable disease untreated.<sup>88</sup> There can be little doubt that the Tuskegee researchers regarded their subjects as less than human.<sup>89</sup> As a result, the ethical canons of experimenting on human subjects were completely disregarded.

The study also raises significant questions about professional self-regulation and scientific bureaucracy. Once the USPHS decided to extend the experiment in the summer of 1933, it was unlikely that the test would be halted short of the men's deaths. The experiment was widely reported for forty years without evoking any significant protest within the medical community. Nor did any bureaucratic mechanism exist within the government for the periodic reassessment of the Tuskegee experiment's ethics and scientific value. The USPHS sent physicians to Tuskegee every several years to check on the study's progress, but never subjected the morality or usefulness of the experiment to serious scrutiny. Only the press accounts of 1972 finally punctured the continued rationalizations of the USPHS and brought the study to an end. Even the HEW investigation was compromised by fear that it would be considered a threat to future human experimentation.

In retrospect the Tuskegee Study revealed more about the pathology of racism than it did about the pathology of syphilis; more about the nature of scientific inquiry than the nature of the disease process. The injustice committed by the experiment went well beyond the facts outlined in the press and the HEW *Final Report*. The degree of deception and damages have been seriously underestimated. As this history of the study suggests, the notion that science is a value-free discipline must be rejected. The need for greater vigilance in assessing the specific ways in which social values and attitudes affect professional behavior is clearly indicated.

## REFERENCES

<sup>1</sup>The best general accounts of the study are “The 40-Year Death Watch,” *Medical World News* (August 18, 1972), pp. 15-17; and

Dolores Katz, "Why 430 Blacks with Syphilis Went Uncured for 40 Years," *Detroit Free Press* (November 5, 1972). The mortality figure is based on a published report of the study which appeared in 1955. See Jesse J. Peters, James H. Peers, Sidney Olansky, John C. Cutler, and Geraldine Gleeson, "Untreated Syphilis in the Male Negro: Pathologic Findings in Syphilitic and Nonsyphilitic Patients," *Journal of Chronic Diseases* 1 (February 1955), 127-48. The article estimated that 30.4 percent of the untreated men would die from syphilitic lesions.

<sup>2</sup>Final Report of the Tuskegee Syphilis Study Ad Hoc Advisory Panel, Department of Health, Education, and Welfare (Washington, D.C.: GPO, 1973). (Hereafter, HEW Final Report).

<sup>3</sup>See George M. Frederickson, *The Black Image in the White Mind* (New York: Harper and Row, 1971), pp. 228-55. Also, John H. Haller, *Outcasts From Evolution* (Urbana, Ill.: University of Illinois Press, 1971), pp. 40-68.

<sup>4</sup>Frederickson, pp. 247-49.

<sup>5</sup>"Deterioration of the American Negro," *Atlanta Journal-Record of Medicine* 5 (July 1903), 287-88. See also J. A. Rodgers, "The Effect of Freedom upon the Psychological Development of the Negro," *Proceedings of the American Medico-Psychological Association* 7 (1900), 88-99. "From the most healthy race in the country forty years ago," concluded Dr. Henry McIlhatton, "he is today the most diseased." "The Sexual Status of the Negro—Past and Present," *American Journal of Dermatology and Genito-Urinary Diseases* 10 (January 1906), 7-9.

<sup>6</sup>W. T. English, "The Negro Problem from the Physician's Point of View," *Atlanta Journal-Record of Medicine* 5 (October 1903), 461. See also, "Racial Anatomical Peculiarities," *New York Medical Journal* 63 (April 1896), 500-01.

<sup>7</sup>"Racial Anatomical Peculiarities," p. 501. Also, Charles S. Bacon, "The Race Problem," *Medicine* (Detroit) 9 (May 1903), 338-43.

<sup>8</sup>H. H. Hazen, "Syphilis in the American Negro," *Journal of the American Medical Association* 63 (August 8, 1914), 463. For deeper background into the historical relationship of racism and sexuality see Winthrop D. Jordan, *White Over Black* (Chapel Hill: University of North Carolina Press, 1968; Pelican Books, 1969), pp. 32-40.

<sup>9</sup>Daniel David Quillian, "Racial Peculiarities: A Cause of the Prevalence of Syphilis in Negroes," *American Journal of Dermatology and Genito-Urinary Diseases* 10 (July 1906), p. 277.

<sup>10</sup>English, p. 463.

<sup>11</sup>William Lee Howard, "The Negro as a Distinct Ethnic Factor in Civilization," *Medicine* (Detroit) 9 (June 1903), 424. See also, Thomas W. Murrell, "Syphilis in the American Negro," *Journal of the American Medical Association* 54 (March 12, 1910), 848.

<sup>12</sup>"Castration Instead of Lynching," *Atlanta Journal-Record of Medicine* 8 (October 1906), 457. The editorial added: "The badge of disgrace and emasculation might be branded upon the face or forehead, as a warning, in the form of an 'R,' emblematic of the crime for which this punishment was and will be inflicted."

<sup>13</sup>Searle Harris, "The Future of the Negro from the Standpoint of the Southern Physician," *Alabama Medical Journal* 14 (January 1902), 62. Other articles on the prevalence of venereal disease among blacks are: H. L. McNeil, "Syphilis in the Southern Negro," *Journal of the American Medical Association* 67 (September 30, 1916), 1001-04; Ernest Philip Boas, "The Relative Prevalence of Syphilis Among Negroes and Whites," *Social Hygiene* 1 (September 1915), 610-16. Doctors went to considerable trouble to distinguish the morbidity and mortality of various diseases among blacks and whites. See, for example, Marion M. Torchia, "Tuberculosis Among American Negroes: Medical Research on a Racial Disease, 1830-1950," *Journal of the History of Medicine and Allied Sciences* 32 (July 1977), 252-79.

<sup>14</sup>Thomas W. Murrell, "Syphilis in the Negro: Its Bearing on the Race Problem," *American Journal of Dermatology and Genito-Urinary Diseases* 10 (August 1906), 307.

<sup>15</sup>"Even among the educated, only a very few will carry out the most elementary instructions as to personal hygiene. One thing you cannot do, and that is to convince the negro that he has a disease that he cannot see or feel. This is due to lack of concentration rather than lack of faith; even if he does believe, he does not care; a child of fancy, the sensations of the passing hour are his only guides to the future." Murrell, "Syphilis in the American Negro," p. 847.

<sup>16</sup>"Deterioration of the American Negro," *Atlanta Journal-Record of Medicine* 5 (July 1903), 288.

<sup>17</sup>Murrell, "Syphilis in the Negro; Its Bearing on the Race Problem," p. 307.

<sup>18</sup>"The anatomical and physiological conditions of the African must be understood, his place in the anthropological scale realized, and his biological basis accepted as being unchangeable by man, before we shall be able to govern his natural uncontrollable sexual passions." See, "As Ye Sow That Shall Ye Also Reap," *Atlanta Journal-Record of Medicine* 1 (June 1899), 266.

<sup>19</sup>Taliaferro Clark, *The Control of Syphilis in Southern Rural Areas* (Chicago: Julius Rosenwald Fund, 1932), 53-58. Approximately 35 percent of the inhabitants of Macon County who were examined were found to be syphilitic.

<sup>20</sup>See Claude Bernard, *An Introduction to the Study of Experimental Medicine* (New York: Dover, 1865, 1957), pp. 5-26.

<sup>21</sup>Taliaferro Clark to M. M. Davis, October 29, 1932. Records of the USPHS Venereal Disease Division, Record Group 90, Box 239, National Archives, Washington National Record Center, Suitland, Maryland. (Hereafter, NA-WNRC). Materials in this collection which relate to the early history of the study were apparently never consulted by the HEW investigation. Included are letters, reports, and memoranda written by the physicians engaged in the study.

<sup>22</sup>H. S. Cumming to R. R. Moton, September 20, 1932, NA-WNRC.

<sup>23</sup>Clark to Davis, October 29, 1932, NA-WNRC.

<sup>24</sup>Cumming to Moton, September 20, 1932, NA-WNRC.

<sup>25</sup>Bruusgaard was able to locate 309 living patients, as well as records from 164 who were deceased. His findings were published as "Ueber das Schicksal der nicht spezifisch behandelten Luetiken," *Archives of Dermatology and Syphilis* 157 (1929), 309-32. The best discussion of the Boeck-Bruusgaard data is E. Gurney Clark and Niels Danbolt, "The Oslo Study of the Natural History of Untreated Syphilis," *Journal of Chronic Diseases* 2 (September 1955), 311-44.

<sup>26</sup>Joseph Earle Moore, *The Modern Treatment of Syphilis* (Baltimore: Charles C. Thomas, 1933), p. 24.

<sup>27</sup>Moore, pp. 231-47; see also John H. Stokes, *Modern Clinical Syphilology* (Philadelphia: W. B. Saunders, 1928), pp. 231-39.

<sup>28</sup>Moore, p. 237.

<sup>29</sup>Moore, p. 236.

<sup>30</sup>J. E. Moore, H. N. Cole, P. A. O'Leary, J. H. Stokes, U. J. Wile, T. Clark, T. Parran, J. H. Usilton, "Cooperative Clinical Studies in the Treatment of Syphilis: Latent Syphilis," *Venereal Disease Information* 13 (September 20, 1932), 351. The authors also concluded that the latently syphilitic were potential carriers of the disease, thus meriting treatment.

<sup>31</sup>Clark to Paul A. O'Leary, September 27, 1932, NA-WNRC. O'Leary, of the Mayo Clinic, misunderstood the design of the study, replying: "The investigation which you are planning in Alabama is indeed an intriguing one, particularly because of the opportunity it affords of observing treatment in a previously untreated group. I assure you such a study is of interest to me, and I shall look forward to its report in the future." O'Leary to Clark, October 3, 1932, NA-WNRC.

<sup>32</sup>Joseph Earle Moore, "Latent Syphilis," unpublished typescript (n.d.), p. 7. American Social Hygiene Association Papers, Social Welfare History Archives Center, University of Minnesota, Minneapolis, Minnesota.

<sup>33</sup>Moore to Clark, September 28, 1932, NA-WNRC. Moore had written in his textbook, "In late syphilis the negro is particularly prone to the development of bone or cardiovascular lesions." See Moore, *The Modern Treatment of Syphilis*, p. 35.

<sup>34</sup>O. C. Wenger to Clark, October 3, 1932, NA-WNRC.

<sup>35</sup>Wenger to Clark, September 29, 1932, NA-WNRC.

<sup>36</sup>Clark Memorandum, September 26, 1932, NA-WNRC. See also, Clark to Davis, October 29, 1932, NA-WNRC.

<sup>37</sup>As Clark wrote: "You will observe that our plan has nothing to do with treatment. It is purely a diagnostic procedure carried out to determine what has happened to the syphilitic Negro who has had no treatment." Clark to Paul A. O'Leary, September 27, 1932, NA-WNRC.


<sup>38</sup>D. G. Gill to O. C. Wenger, October 10, 1932, NA-WNRC.



- <sup>39</sup>Clark to Vonderlehr, January 25, 1933, NA-WNRC.
- <sup>40</sup>Vonderlehr to Clark, February 28, 1933, NA-WNRC.
- <sup>41</sup>Vonderlehr to Clark, November 2, 1932, NA-WNRC. Also, Vonderlehr to Clark, February 6, 1933, NA-WNRC.
- <sup>42</sup>Clark to Vonderlehr, March 9, 1933, NA-WNRC.
- <sup>43</sup>Vonderlehr later explained: "The reason treatment was given to many of these men was twofold: First, when the study was started in the fall of 1932, no plans had been made for its continuation and a few of the patients were treated before we fully realized the need for continuing the project on a permanent basis. Second it was difficult to hold the interest of the group of Negroes in Macon County unless some treatment was given." Vonderlehr to Austin V. Diebert, December 5, 1938, Tuskegee Syphilis Study Ad Hoc Advisory Panel Papers, Box 1, National Library of Medicine, Bethesda, Maryland. (Hereafter, TSS-NLM). This collection contains the materials assembled by the HEW investigation in 1972.
- <sup>44</sup>Vonderlehr to Clark, February 6, 1933, NA-WNRC.
- <sup>45</sup>H. S. Cumming to J. N. Baker, August 5, 1933, NA-WNRC.
- <sup>46</sup>January 22, 1933; January 12, 1933, NA-WNRC.
- <sup>47</sup>Vonderlehr to Clark, January 28, 1933, NA-WNRC.
- <sup>48</sup>Clark to Moore, March 25, 1933, NA-WNRC.
- <sup>49</sup>Macon County Health Department, "Letter to Subjects," n.d., NA-WNRC.
- <sup>50</sup>Vonderlehr to Clark, April 8, 1933, NA-WNRC. See also, Vonderlehr to Wenger, July 18, 1933, NA-WNRC.
- <sup>51</sup>Wenger to Vonderlehr, July 21, 1933, NA-WNRC. The italics are Wenger's.
- <sup>52</sup>Cumming to Moton, July 27, 1933, NA-WNRC.
- <sup>53</sup>Wenger to Vonderlehr, July 21, 1933, NA-WNRC.
- <sup>54</sup>Vonderlehr to Murray Smith, July 27, 1933, NA-WNRC.
- <sup>55</sup>Wenger to Vonderlehr, August 5, 1933, NA-WNRC.
- <sup>56</sup>Vonderlehr to Wenger, October 24, 1933, NA-WNRC. Controls were given salicylates.
- <sup>57</sup>Austin V. Diebert and Martha C. Bruyere, "Untreated Syphilis in the Male Negro, III," *Venereal Disease Information* 27 (December 1946), 301-14.
- <sup>58</sup>Eunice Rivers, Stanley Schuman, Lloyd Simpson, Sidney Olansky, "Twenty-Years of Followup Experience In a Long-Range Medical Study," *Public Health Reports* 68 (April 1953), 391-95. In this article Nurse Rivers explains her role in the experiment. She wrote: "Because of the low educational status of the majority of the patients, it was impossible to appeal to them from a purely scientific approach. Therefore, various methods were used to maintain their interest. Free medicines, burial assistance or insurance (the project being referred to as 'Miss Rivers' Lodge'), free hot meals on the days of examination, transportation to and from the hospital, and an opportunity to stop in town on the return trip to shop or visit with their friends on the streets all helped. In spite of these attractions, there were some who refused their examinations because they were not sick and did not see that they were being benefitted." (p. 393).
- <sup>59</sup>Austin V. Diebert to Raymond Vonderlehr, March 20, 1939, TSS-NLM, Box 1.
- <sup>60</sup>Murray Smith to Subjects, (1938), TSS-NLM, Box 1. See also, Sidney Olansky to John C. Cutler, November 6, 1951, TSS-NLM, Box 2.
- <sup>61</sup>The USPHS originally requested that the Julius Rosenwald Fund meet this expense. See Cumming to Davis, October 4, 1934, NA-WNRC. This money was usually divided between the undertaker, pathologist, and hospital. Lloyd Isaacs to Raymond Vonderlehr, April 23, 1940, TSS-NLM, Box 1.
- <sup>62</sup>Stanley H. Schuman, Sidney Olansky, Eunice Rivers, C. A. Smith, Dorothy S. Rambo, "Untreated Syphilis in the Male Negro: Background and Current Status of Patients in the Tuskegee Study," *Journal of Chronic Diseases* 2 (November 1955), 555.
- <sup>63</sup>R. A. Vonderlehr and Taliaferro Clark, "Untreated Syphilis in the Male Negro," *Venereal Disease Information* 17 (September 1936), 262.
- <sup>64</sup>J. R. Heller and P. T. Bruyere, "Untreated Syphilis in the Male Negro: II. Mortality During 12 Years of Observation," *Venereal Disease Information* 27 (February 1946), 34-38.
- <sup>65</sup>Jesse J. Peters, James H. Peers, Sidney Olansky, John C. Cutler, and Geraldine Gleeson, "Untreated Syphilis in the Male Negro: Pathologic Findings in Syphilitic and Non-Syphilitic Patients," *Journal of Chronic Diseases* 1 (February 1955), 127-48.
- <sup>66</sup>Sidney Olansky, Stanley H. Schuman, Jesse J. Peters, C. A. Smith, and Dorothy S. Rambo, "Untreated Syphilis in the Male Negro, X. Twenty Years of Clinical Observation of Untreated Syphilitic and Presumably Nonsyphilitic Groups," *Journal of Chronic Diseases* 4 (August 1956), 184.
- <sup>67</sup>O. C. Wenger, "Untreated Syphilis in Male Negro," unpublished typescript, 1950, p. 3. Tuskegee Files, Center for Disease Control, Atlanta, Georgia. (Hereafter TF-CDC).
- <sup>68</sup>Vernal G. Cave, "Proper Uses and Abuses of the Health Care Delivery System for Minorities with Special Reference to the Tuskegee Syphilis Study," *Journal of the National Medical Association* 67 (January 1975), 83.
- <sup>69</sup>See for example, Vonderlehr to B. W. Booth, April 18, 1934; Vonderlehr to E. R. Lett, November 20, 1933, NA-WNRC.
- <sup>70</sup>"Transcript of Proceedings—Tuskegee Syphilis Ad Hoc Advisory Panel," February 23, 1973, unpublished typescript, TSS-NLM, Box 1.
- <sup>71</sup>Raymond Vonderlehr to Murray Smith, April 30, 1942; and Smith to Vonderlehr, June 8, 1942, TSS-NLM, Box 1.
- <sup>72</sup>Stanley H. Schuman, Sidney Olansky, Eunice Rivers, C. A. Smith, and Dorothy S. Rambo, "Untreated Syphilis in the Male Negro: Background and Current Status of Patients in the Tuskegee Study," *Journal of Chronic Diseases* 2 (November 1955), 550-53.
- <sup>73</sup>Raymond Vonderlehr to Stanley H. Schuman, February 5, 1952, TSS-NLM, Box 2.
- <sup>74</sup>Schuman et al., p. 550.
- <sup>75</sup>"Minutes, April 5, 1965" unpublished typescript, TSS-NLM, Box 1.
- <sup>76</sup>"Tuskegee Ad Hoc Committee Meeting—Minutes, February 6, 1969," TF-CDC.
- <sup>77</sup>James B. Lucas to William J. Brown, September 10, 1970, TF-CDC.
- <sup>78</sup>Elizabeth M. Kennebrew to Arnold C. Schroeter, February 24, 1971, TSS-NLM, Box 1.
- <sup>79</sup>See *Medical Tribune* (September 13, 1972), pp. 1, 20; and Report on HEW's Tuskegee Report, *Medical World News* (September 14, 1973), pp. 57-58.
- <sup>80</sup>HEW Final Report, p. 7.
- <sup>81</sup>The notable exception is Jay Katz's eloquent "Reservations About the Panel Report on Charge 1," HEW Final Report, pp. 14-15.
- <sup>82</sup>HEW Final Report, pp. 8, 12.
- <sup>83</sup>HEW Final Report, pp. 8, 12.
- <sup>84</sup>See R. H. Kampmeier, "The Tuskegee Study of Untreated Syphilis," *Southern Medical Journal* 65 (October 1972), 1247-51; and "Final Report on the 'Tuskegee Syphilis Study,'" *Southern Medical Journal* 67 (November 1974), 1349-53.
- <sup>85</sup>Leonard J. Goldwater, "The Tuskegee Study in Historical Perspective," unpublished typescript, TSS-NLM; see also "Treponemes and Tuskegee," *Lancet* (June 23, 1973), p. 1438; and Louis Lasagna, *The VD Epidemic* (Philadelphia: Temple University Press, 1975), pp. 64-66.
- <sup>86</sup>Quoted in "Debate Revives on the PHS Study," *Medical World News* (April 19, 1974), p. 37.
- <sup>87</sup>Heller to Vonderlehr, November 28, 1933, NA-WNRC; quoted in *Medical Tribune* (August 23, 1972), p. 14.
- <sup>88</sup>Although it is now known that syphilis is rarely infectious after its early phase, at the time of the study's inception latent syphilis was thought to be communicable. The fact that members of the control group were placed in the test group when they became syphilitic proves that at least some infectious men were denied treatment.
- <sup>89</sup>When the subjects are drawn from minority groups, especially those with which the researcher cannot identify, basic human rights may be compromised. Hans Jonas has clearly explicated the problem in his "Philosophical Reflections on Experimentation," *Daedalus* 98 (Spring 1969), 234-37. As Jonas writes: "If the properties we adduced as the particular qualifications of the members of the scientific fraternity itself are taken as general criteria of selection, then one should look for additional subjects where a maximum of identification, understanding, and spontaneity can be expected—that is, among the most highly motivated, the most highly educated, and the least 'captive' members of the community."



# A New Look at an Old Problem: Improving Diversity, Equity, and Inclusion in Scientific Research

The American Surgeon  
2021, Vol. 87(11) 1722–1726  
© The Author(s) 2021  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/00031348211029853  
[journals.sagepub.com/home/asu](https://journals.sagepub.com/home/asu)  


**Brandon V. Henry, MD, MPH<sup>1</sup>, Herbert Chen, MD<sup>2</sup>,  
Michael A. Edwards, MD<sup>3</sup>, Lydia Faber, BS<sup>4</sup>, and Julie A.  
Freischlag, MD<sup>4</sup>**

## Introduction

Almost 30 years ago, Congress passed the National Institutes of Health (NIH) Revitalization Act of 1993 which required that women and members of minority groups be included in all NIH-funded clinical research.<sup>1</sup> This act was created “to ensure that all NIH-funded research will be carried out in a manner sufficient to elicit information about individuals of both sexes/genders and diverse racial and ethnic groups...”<sup>2</sup> Recent years have seen many firsts for women and minorities, most notably the first female and woman of color becoming vice president of the United States. In addition, the importance of diversity, equity, and inclusion (DEI) has become an increasingly prevalent part of social and political discourse. However, women and racial/ethnic minorities remain underrepresented in basic science, clinical trials, and publications. Research funding inequities and underrepresentation of these groups in research-related peer review organizations and processes compound this problem.

Research disparities that limit the generalizability of published data are likely to widen racial, ethnic, and gender treatment gaps and health inequities. Certainly scientific research on the intersectionality of structural racism, cultural racism, and discrimination as a fundamental cause of racial inequalities in health has increased.<sup>3</sup> However, under-representation in both our health care and research enterprises contribute to inequity in scientific inquiry,<sup>4</sup> and ultimately can perpetuate gender, sex-related, and racial health inequities. Developing and implementing effective mitigation strategies will require exploration of these gaps in the context of racialized societal structures and sex-biased norms that result in inequities in research design, funding, peer review, and editorial oversight.

## Bias in Basic Science and Clinical Trial Research

Sex bias heavily impacts science, from the basic sciences lab to the editorial boards that oversee publishing. The

roots of sex bias in basic science research can be traced to as early as 1923 when an experiment attributed locomotor differences between male and female rats to the estrous cycle.<sup>5</sup> Despite this conclusion being unfounded,<sup>5</sup> researchers have continued to utilize male-only subjects due to the possibility that the estrous cycle could impact the analysis of non-reproductive mechanisms.<sup>6</sup> Other reasons for using male-only subjects include cost concerns, inconvenience, and unconscious bias against the importance of sex differences.<sup>6</sup>

This problem in academic medical research stems from limitations in both sex reporting and animal selection. One study reviewed 2347 articles across five preeminent surgical journals (2011–2012), finding that 80% of the publications that specified sex for animal research reported male animals only, 17% female only, and 3% analyzed both sexes.<sup>2</sup> Overall, 22% of the publications did not specify sex at all.<sup>6</sup> The number of animals tested was also disproportionate at 84% male and 16% female.<sup>6</sup> Another study found that only 31% of biomedical research papers report sex.<sup>7</sup>

In clinical trials, the sex ratio is closer to equal, averaging 51.1% male and 49.0% female.<sup>8</sup> However, it is not equitable when analyzed based on trial phase, funding source, and randomization. The largest gap is observed in Phase I clinical trials, where 64.1% of the subjects are male and 35.9% are female.<sup>8</sup> The gender gap decreases in Phase 2 and 3 trials to 48.4% and 51.0% male, respectively.<sup>8</sup> This exclusion in early stages is highly problematic in the critical time when drug dosing and

<sup>1</sup>Department of Surgery, Morehouse School of Medicine, Atlanta, GA, USA

<sup>2</sup>Department of Surgery, University of Alabama at Birmingham, Birmingham, AL, USA

<sup>3</sup>Department of Surgery, Mayo Clinic Hospital, Jacksonville, FL, USA

<sup>4</sup>Department of Surgery, Wake Forest School of Medicine, Winston-Salem, NC, USA

### Corresponding Author:

Brandon V. Henry MD, Department of Surgery, Morehouse School of Medicine, 720 Westview Drive, SW, Atlanta, GA 30310, USA.  
Email: [bvhenry@msm.edu](mailto:bvhenry@msm.edu)

safety occur. The U.S. Government Accountability Office reported that eight of ten medications recalled by the FDA were due to different reactions seen in males and females.<sup>9</sup>

A similar problem exists in clinical trials concerning ethnicity and race. In reviewing 270 published RCTs, Rochon et al found that 70% had no race or ethnicity data. Non-white participation rate in the remaining third was 12-13%.<sup>10</sup> A more recent review revealed no change in inclusion trends.<sup>11</sup> These trends are reported across a wide spectrum of medical literature. Massie et al<sup>12</sup> found racial inequity in published color photographs, relative to population representation. Cho et al<sup>13</sup> found only 6-8% of breast surgery images depicted a non-white patient. Addressing such disparities will be essential to eliminating bias in constructing research projects in the future.

## Inequities in Academic Advancement and Peer Review

Problems related to DEI in scientific research certainly go beyond the known biases seen in the structure and design of basic science projects and clinical trials. One of the important outputs of research is the dissemination of information and discovery that most often occurs through publication in scientific journals. The process of medical journal publication involves peer review of the data. Participation in this peer review process is critical to academic development, including supporting extramural fundability. Disparities in representation at this level can have significant impacts, not only in the form of publication bias, but also in biasing the path to academic advancement. The cyclical effect of this problem can be demonstrated by the underrepresentation of women in positions of leadership in academic medicine.

Currently, women comprise a majority of medical school matriculants at 50.5%, but leadership positions in academic medicine have not kept pace.<sup>14</sup> In a 2015 study, 38% of full-time faculty positions were held by women, and just 21% of women held full professorships.<sup>15</sup> In addition, men are twice as likely to receive an invitation from a journal to submit their research, and more than two-thirds of grant applications are from male investigators.<sup>16</sup> In reviewing over 5.5 million publications and 27 million authorships across various disciplines, West et al reported over 70% and 66% were authored and first-authored by men.<sup>17</sup> Each of these factors increases male chances to achieve full professorships,<sup>16</sup> and due to their greater representation among faculty, they are more likely to be elected to medical journal editorial boards.

In a review of 60 top medical journals, only 15.9% of the editors-in-chief were female, and 17.5% of editorial board members were women.<sup>18</sup> Women are also greatly underrepresented in six major medical journals (*JAMA*, *JAMA Internal Medicine*, *The British Medical Journal*,

*The Lancet*, *The New England Journal of Medicine*, and *Annals of Internal Medicine*) as the percentage of female board reviewers ranged between 16.6-28.8% in 2010-2011.<sup>19</sup> Even in women's health journals, females are the minority, as 42% of the 1440 board members are women, and 41% of editor-in-chief positions are held by women.<sup>16</sup> Shockingly, females comprise 61.9% of pediatrics, yet none of the top four pediatric journals had a female editor in chief as of 2018.<sup>16</sup>

While the gender gap in the peer review processes has been well-established,<sup>17,20-22</sup> less has been published on racial and ethnic peer review gaps. There is increasing recognition that medical journals have a critical role to play in addressing peer review inequity, and ultimately reducing racial and ethnic health inequities<sup>23,24</sup>; however, this remains aspirational. In a recent non-peer reviewed opinion article, the author reported editorial staff racial and ethnic diversity for the NEJM and JAMA to be 3.9% and 8.2%, respectively.<sup>25</sup> Similar to women in academic medicine, the road to advancement for minorities may also be limited by a lack of representation in the peer-review process.

## Disparities in Research Funding

Funding disparities also play a role as funded research often informs clinical practice patterns and health policy.<sup>26,27</sup> Therefore, funding priorities can dictate societal priorities and policies, which can either mitigate or perpetuate health inequities.<sup>26</sup> In the early 2000s, several initiatives, including establishing the National Institute on Minority Health and Health Disparities (NIMHD), were catalysts for health disparities research. This resulted in increased funding for research in the domains of health disparities, health inequities, and social determinants of health (SDOH). However, disparities in funding persist even within these research domains.

Kneipp et al<sup>28</sup> explored funding patterns by NIH institutes to determine if funding patterns were aligned with broadening understanding of the relationship between health inequity/SDOH and racial/ethnic inequalities in health outcomes. Over 17-years, 32 968 projects were funded. Funding for health disparity projects was 14 to 19-fold higher than that of funding for inequity and SDOH research. Compared to the NIMHD, funding of health disparity, health inequity, or SDOH research by other NIH institutes was 2 to 30-fold lower. In addition, across institutes health inequity/SDOH research funding rate was .2% to 8.9%,<sup>28</sup> highlighting the gap between identified research focus needs and actual funding allocation.

Under funding of women, racial/ethnic minority investigators, and public institutions is also well-established.<sup>29</sup> Even though NIH awards have improved for these investigator cohorts, there remain significant funding gaps.<sup>30</sup> While the presence of these persistent

gaps is multifactorial, the role of entrenched biases in funding processes cannot be discounted.<sup>31</sup> Several potential trickle-down effects can occur in the presence of persistent funding disparities. These include decreased diversity in the medical research workforce, less research participation within minority groups, and widening of the research prioritization gap.<sup>32-35</sup> Biases, stereotypes, and mistrust of the research workforce are impediments to recruitment diversity in clinical trials and subsequent representation in publications.<sup>36,37</sup> Solving the problem of disparate funding is just as important to creating equitable research as addressing biases in publication.

## Improving Editorial Board DEI

Currently, there is a push for equity and inclusion amongst editorial board members.<sup>38</sup> In response to a recent controversial paper published in the *Journal of Vascular Surgery*, there was a call for increased gender diversity in the review and publication process. The editors of 108 surgical journals signed on to a Call to Action<sup>39</sup>: “The current and former Editors in Chief of *JAMA Surgery*, announce this call to action for all surgery journal editors. *JAMA Surgery* hereby publicly commits to diversity and inclusion throughout the editorial and peer review process. In addition, the journal commits to improving diversity among the editorial leadership and editorial board of the journal and when inviting authors to write commentaries for the journal. We ask that all surgery journal editors make a pledge to strive for diversity in the peer review process and among the editorial boards and editorial leaders.”

Additionally, in June 2020, in response to the events that highlighted racism, hatred, and social injustices as well as the disparities in health care brought by COVID-19, the Editors of the *American Journal of Surgery* released a statement calling for the following action<sup>40</sup>: “(1) Ensuring our editorial board has diverse representation to assure all voices are heard and that we are addressing issues of racism, discrimination and structural violence in an open and inclusive way. (2) Creating a space which values and encourages rigorous research and scholarship into the areas of health inequities and will move the field forward. We will put out a call for papers in this domain and make this content free for all to view, (3) Using our resources and influence to support charities which address health inequities and racial discrimination.”

There remains, however, a lack of data on DEI in the publication process. Is there discrimination in the publication process? Does the composition of the surgical journal editorial board effect acceptance rate for different papers? What is the current composition of editorial boards for surgical journals? Does the acceptance rate of DEI related manuscripts vary between surgical journals? To answer these questions, Erin White and Gurit Sandhu

have recently conduct a survey of the editorial boards of 6 surgical journals, *American Journal of Surgery*, *Annals of Surgery*, *Surgery*, *Journal of Surgical Research*, *Journal of Surgical Education*, and *JAMA Surgery*. Of the 240 board members, 50% responded. They found that there were gender, racial, and ethnic differences between the 6 journal boards.<sup>41</sup> *American Journal of Surgery* had the highest gender diversity with 44% of the board comprised of women. In the second phase of the study, 5 of the above journals have agreed to provide access to the papers published between 2016 to present to determine if the acceptance rate differs for DEI related manuscripts. This second phase is still in progress. However, early data suggested that DEI related manuscripts have a higher rate of acceptance compared to other articles in the *American Journal of Surgery*. Therefore, early data suggest that there is much more work to do with regards to increasing diversity on editorial boards of surgical journals.

## Technology and the Future of DEI in Scientific Research

As demonstrated above, improving DEI in surgical research is a multifaceted challenge. Addressing the necessity of diverse representation among researchers, the challenges in structuring equitable research projects, the importance of identifying and mitigating potential biases in research funding allocation, and the value of increased diversity among editorial boards will require significant intentional effort by the surgical research community. Accomplishing these goals will require deep introspection and inquiry regarding the structural mechanisms in place that have created disparities in our field. Objective data driven methods will be necessary in this process lest our best intentions be hampered by unappreciated subjective biases.

Thus, we should consider doing everything we can to mitigate these biases. Utilizing technology may provide a way to accomplish this. For example, name bias has been demonstrated to have potentially discriminatory effects in multiple settings including hiring practices and in the classroom.<sup>42-44</sup> Computer software can be used to blind reviewers of grant applications and publication submissions to help decrease the possibility of such biases. Even blinding of institutional names can proactively help to mitigate any subconscious biases that could affect submissions associated with institutions that are known to hire and serve underrepresented groups.

It may also prove helpful to look to other industries for strategies that have successfully improved DEI in the workplace. Companies in the technology sector have employed analytics to tackle this problem with some success over the past decade. Many have increasingly turned to artificial intelligence (AI) and data driven approaches to improving DEI. For example, writing analysis

software has been used to reduce bias in the language of job postings, thus creating more gender-neutral text. This can potentially attract a more balanced applicant pool.<sup>45</sup> Furthermore, analytic tools can be used to identify weakness in recruitment and retention efforts in these environments.<sup>46</sup> Google has utilized internal human resources data and external data from the U.S. Equal Employment Opportunity Commission to identify areas of weakness and to create mathematical algorithms to help improve DEI focused hiring and retention practices. These data-driven efforts have yielded positive results. In the company's 2020 Diversity Annual Report, the company stated that due to these strategies in the previous year, they achieved the largest increase in underrepresented minority hires the company has ever recorded, and have they seen an 11% increase in company applications from women.<sup>47</sup>

While even these technologies are not perfect and can be subject to bias in how their algorithms are constructed, we should not shy away from employing such sophisticated and objective tools in our efforts to create a more equitable surgical research community. Increasing diversity and representation at all levels, from the recruitment of researchers, to the design of our projects, and to the construction of our editorial boards should be imperative. This will help to enrich not only the personal and collective experiences of members of our community, but also help to improve the context for the data and results we generate now and in the future.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### References

- Heller C, Balls-Berry JE, Nery JD, Erwin PJ, Littleton D, Kim M, et al. Strategies addressing barriers to clinical trial enrollment of underrepresented populations: A systematic review. *Contemp Clin Trials*. 2014;39(2):169-182.
- National Institutes of Health. *NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research*. NIH grants and funding. <https://grants.nih.gov/policy/inclusion/women-and-minorities/guidelines.htm> (accessed 25 May 2021).
- Williams DR, Lawrence JA, Davis BA. Racism and health: Evidence and needed research. *Annu Rev Publ Health*. 2019;40:105-125.
- Nature. Henrietta Lacks: Science must right a historical wrong. *Nature*. 2020;585(7823):7.
- Prendergast BJ, Onishi KG, Zucker I. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci Biobehav Rev*. 2014;40:1-5.
- Yoon DY, Mansukhani NA, Stubbs VC, Helenowski IB, Woodruff TK, Kibbe MR. Sex bias exists in basic science and translational surgical research. *Surgery*. 2014;156(3):508-516. doi: [10.1016/j.surg.2014.07.001](https://doi.org/10.1016/j.surg.2014.07.001)
- Sugimoto CR, Ahn Y-Y, Smith E, Macaluso B, Larivière V. Factors affecting sex-related reporting in medical research: A cross-disciplinary bibliometric analysis. *Lancet*. 2019;393(10171):550-559.
- Prakash VS, Mansukhani NA, Helenowski IB, Woodruff TK, Kibbe MR. Sex bias in interventional clinical trials. *J Wom Health*. 2018;27(11):1342-1348. doi: [10.1089/jwh.2017.6873](https://doi.org/10.1089/jwh.2017.6873)
- Mansukhani NA, Yoon DY, Teter KA, Stubbs VC, Helenowski IB, Woodruff TK, et al. Determining if sex bias exists in human surgical clinical research. *JAMA Surg*. 2016;151(11):1022-1030.
- Rochon, PA, Mashari A, Cohen A, Misra A, Laxer D, Streiner DL, et al. The inclusion of minority groups in clinical trials: Problems of under representation and under reporting of data. *Account Res*. 2004;11(3-4):215-223.
- Geller SE, Koch A, Pellettieri B, Carnes M. Inclusion, analysis, and reporting of sex and race/ethnicity in clinical trials: Have we made progress? *J Wom Health*. 2011;20(3):315-320.
- Massie JP, Cho DY, Kneib CJ, Burns JR, Crowe CS, Lane M, et al. Patient representation in medical literature: Are we appropriately depicting diversity? *Plast Reconstr Surg Glob Open*. 2019;7(12):e2563.
- Cho DY, Kneib CJ, Shakir A, Burns JR, Lane M, Massie JP, et al. Underrepresentation of racial minorities in breast surgery literature: A call for increased diversity and inclusion. *Ann Surg*. 2021;273(2):202-207.
- Heiser S. *The majority of U.S. medical students are women, new data show*. AAMC. <https://www.aamc.org/news-insights/press-releases/majority-us-medical-students-are-women-new-data-show> (Accessed April 30, 2021).
- Jena AB, Khullar D, Ho O, Olenski AR, Blumenthal DM. Sex differences in academic rank in US medical schools in 2014. *J Am Med Assoc*. 2015;314(11):1149-1158.
- Grinnell M, Higgins S, Yost K, Ochuba O, Lobl M, Grimes P, et al. The proportion of male and female editors in women's health journals: A critical analysis and review of the sex gap. *Int J Womens Dermatol*. 2020;6(1):7-12.
- Amrein K, Langmann A, Fahrleitner-Pammer A, Pieber TR, Zollner-Schwetz I. Women underrepresented on editorial boards of 60 major medical journals. *Gend Med*. 2011;8(6):378-387.
- Erren TC, Groß JV, Shaw DM, Selle B. Representation of women as authors, reviewers, editors in chief, and editorial board members at 6 general medical journals in 2010 and 2011. *JAMA Intern Med*. 2014;174(4):633-635.



19. Jaggi, R, Guancial EA, Worobey CC, Henault LE, Chang Y, Starr R, et al. The “gender gap” in authorship of academic medical literature—A 35-year perspective. *N Engl J Med*. 2006;355(3):281-287.
20. Anonymous. Nature’s under-representation of women. *Nature*. 2018;558(7710):344.
21. McGinty GB. Building diversity one citation at a time. *J Am Coll Radiol*. 2018;15(7):937-938.
22. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. The role of gender in scholarly authorship. *PLoS One*. 2013; 8(7):e66212.
23. Shah SS, Shaughnessy EE, Spector ND. Leading by example: How medical journals can improve representation in academic medicine. *J Hosp Med*. 2019;14(7):393.
24. Elston DM. Medical journals, diversity, and cultural sensitivity. *J Am Acad Dermatol*. 2021;84(3):607-608.
25. McFarling UL. *When a cardiologist flagged the lack of diversity at premier medical journals, the silence was telling*. STAT, <https://www.statnews.com/2021/04/12/lack-of-diversity-at-premier-medical-journals-jama-nejm/> (Accessed May 29, 2021).
26. MacKay K, Quigley M. Exacerbating inequalities? Health policy and the behavioural sciences. *Health Care Anal*. 2018;26(4):380-397.
27. Smits PA, Denis JL. How research funding agencies support science integration into policy and practice: An international overview. *Implement Sci*. 2014;9:28.
28. Kneipp SM, Schwartz TA, Drevdahl DJ, Canales MK, Santacroce S, Santos HP, et al. Trends in health disparities, health inequity, and social determinants of health research: A 17-year analysis of NINR, NCI, NHLBI, and NIMHD funding. *Nurs Res*. 2018;67(3):231-241.
29. Mesquita-Neto JW, Dailey W, Macedo FI, Merchant NB. Patterns of national institutes of health grant funding to surgical research and scholarly productivity in the United States. *Ann Surg*. 2020;272(4):539-546.
30. Nikaj S, Roychowdhury D, Lund PK, Matthews M, Pearson K. Examining trends in the diversity of the U.S. national institutes of health participating and funded workforce. *FASEB J*. 2018;32(18):639.
31. Kaatz A, Magua W, Zimmerman DR, Carnes M. A quantitative linguistic analysis of national institutes of health R01 application critiques from investigators at one institution. *Acad Med*. 2015;90(1):69-75.
32. Chen MS Jr. Rectifying disparities in funding of Asian American, native Hawaiian, and Pacific Islander research by the US national institutes of health. *JAMA Netw Open*. 2019;2(7):e197561.
33. Durant RW, Wenzel JA, Scarinci IC, Paterniti DA, Fouad MN, Hurd TC, et al. Perspectives on barriers and facilitators to minority recruitment for clinical trials among cancer center leaders, investigators, research staff, and referring clinicians: Enhancing minority participation in clinical trials (EMPaCT). *Cancer*. 2014;120(Suppl 7):1097-1105.
34. Robinson JM, Trochim WM. An examination of community members’, researcher’ and health professionals’ perceptions of barriers to minority participation in medical research: An application of concept mapping. *Ethn Health*. 2007;12(5):521-539.
35. Amorrortu RP, Arevalo M, Vernon SW, Mainous AG, Diaz V, McKee MD, et al. Recruitment of racial and ethnic minorities to clinical trials conducted within specialty clinics: An intervention mapping approach. *Trials*. 2018; 19(1):115.
36. Niranjani SJ, Martin MY, Fouad MN, Vickers SM, Wenzel JA, Cook ED, et al. Bias and stereotyping among research and clinical professionals: Perspectives on minority recruitment for oncology clinical trials. *Cancer*. 2020;126(9): 1958-1968.
37. Tilley BC, Mainous III AG, Smith DW, McKee MD, Amorrortu RP, Alvidrez J, et al. Design of a cluster-randomized minority recruitment trial: RECRUIT. *Clin Trials*. 2017; 14(3):286-298.
38. Mehta A, Morris-Wiseman L, Romero Arenas MA, Nwariaku F. Beyond recognition: practical steps to inclusion in academic surgery. *Am J Surg*. 2021:S0002-9610(21)00156-2.
39. Kibbe MR, Freischlag J. Call to action to all surgery journal editors for diversity in the editorial and peer review process. *JAMA Surgery*. 2020;155(11):1015-1016. doi:10.1001/jamasurg.2020.4549.
40. Chen H, Ferrada P, Kennedy G, Lindeman B, Newman E, Solorzano C, et al. Editorial: From the editors of AJS. *Am J Surg* 2020;220:507.
41. White E, Maduka R, Ballouz D, et al. Surgical research journals - under review. an assessment of diversity among editorial boards and outcomes of peer review. *Am J Surg* (in press).
42. Cotton JL, O’Neill BS, Griffin A. The “name game”: Affective and hiring reactions to first names. *J Manag Psychol*. 2008;23(1):18-39.
43. Conaway W, Bethune S. Implicit bias and first name stereotypes: What are the implications for online instruction? *Online Learn*. 2015;19(3):162-178.
44. MacNeill L, Driscoll A, Hunt AN. What’s in a name: Exposing gender bias in student ratings of teaching. *Innovat High Educ*. 2015;40:291–303.
45. Ansari K. *Learn how to remove gender bias from job ads*. Undercover Recruiter. <https://theundercoverrecruiter.com/remove-biased-job-ads/> (Accessed April 29, 2021).
46. Sullivan J. *How google is using people analytics to completely reinvent HR*, TLNT. <https://www.tlnt.com/how-google-is-using-people-analytics-to-completely-reinvent-hr/> (Accessed April 29, 2021).
47. *Google diversity annual report 2020*. Google. <https://diversity.google/annual-report/> (Accessed April 30, 2021).

## Historical and Philosophical Perspective

# The Mismeasure of Science: Stephen Jay Gould versus Samuel George Morton on Skulls and Bias

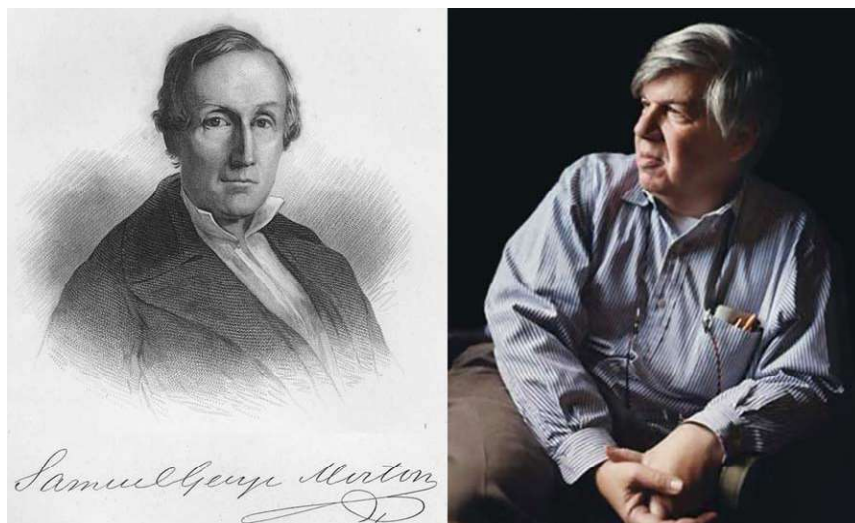
Jason E. Lewis<sup>1\*</sup>, David DeGusta<sup>2</sup>, Marc R. Meyer<sup>3</sup>, Janet M. Monge<sup>4</sup>, Alan E. Mann<sup>5</sup>, Ralph L. Holloway<sup>6</sup>

**1** Department of Anthropology, Stanford University, Stanford, California, United States of America, **2** Paleoanthropology Institute, Oakland, California, United States of America, **3** Department of Anthropology, Chaffey College, Rancho Cucamonga, California, United States of America, **4** Department of Anthropology and Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **5** Department of Anthropology, Princeton University, Princeton, New Jersey, United States of America, **6** Department of Anthropology, Columbia University, New York, New York, United States of America

Stephen Jay Gould, the prominent evolutionary biologist and science historian, argued that “unconscious manipulation of data may be a scientific norm” because “scientists are human beings rooted in cultural contexts, not automatons directed toward external truth” [1], a view now popular in social studies of science [2–4]. In support of his argument Gould presented the case of Samuel George Morton, a 19th-century physician and physical anthropologist famous for his measurements of human skulls. Morton was considered the objectivist of his era, but Gould reanalyzed Morton’s data and in his prize-winning book *The Mismeasure of Man* [5] argued that Morton skewed his data to fit his preconceptions about human variation. Morton is now viewed as a canonical example of scientific misconduct. But did Morton really fudge his data? Are studies of human variation inevitably biased, as per Gould, or are objective accounts attainable, as Morton attempted? We investigated these questions by remeasuring Morton’s skulls and reexamining both Morton’s and Gould’s analyses. Our results resolve this historical controversy, demonstrating that Morton did not manipulate data to support his preconceptions, contra Gould. In fact, the Morton case provides an example of how the scientific method can shield results from cultural biases.

## A Debate across a Century

Stephen Jay Gould (1941–2002) and Samuel George Morton (1799–1851) worked in different centuries but shared a number of similarities (Figure 1). Each was well-known to the public and held a prominent academic position, Morton as president of Philadelphia’s Academy of Natural Sciences, Gould as a Harvard professor. Gould’s popular books on science were best-sellers, and Morton’s 1839 *Crania Americana* volume brought him international renown. Both had an exceptionally broad range of research interests that included invertebrate paleontology—Morton was the first American practitioner of this discipline [6], and it was with



**Figure 1. Portraits of Samuel George Morton (left) and Stephen Jay Gould (right).** Morton’s portrait is from [22] and Gould’s photo was taken by Kathy Chapman and is used under a Creative Commons Attribution 3.0 license. doi:10.1371/journal.pbio.1001071.g001

studies of fossil land snails that Gould initially made his mark [7]. But it was Morton’s work on human skulls that drew first Gould’s interest, then his ire.

Much of Morton’s fame derived from his “American Golgotha”—a collection of nearly 1,000 human skulls (Figure 2) he obtained from around the world [6]. Morton took detailed measurements of these skulls with a particular focus on cranial capacity, the skeletal equivalent of brain size [8–10]. From these measurements he hoped to determine

whether different human populations were separate species resulting from multiple divine creations (polygenesis) or a single species created but once (monogenesis), a major question in pre-Darwinian science [6]. Morton’s empirical approach, generating data by systematically measuring large numbers of actual specimens, was groundbreaking and he was considered the objectivist of his era [1,6]. Even so, as the polygenesis-monogenesis debate faded, Morton’s work lapsed into relative obscurity [6,11].

**Citation:** Lewis JE, DeGusta D, Meyer MR, Monge JM, Mann AE, et al. (2011) The Mismeasure of Science: Stephen Jay Gould versus Samuel George Morton on Skulls and Bias. PLoS Biol 9(6): e1001071. doi:10.1371/journal.pbio.1001071

**Published:** June 7, 2011

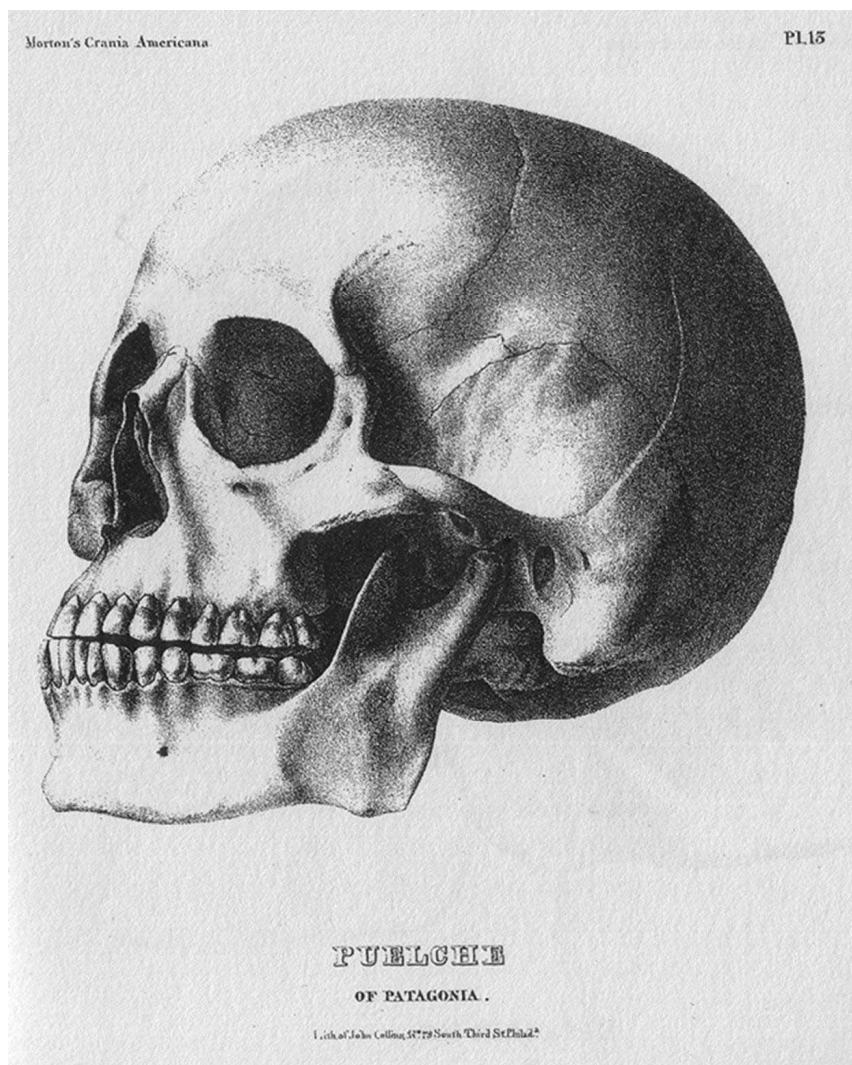
**Copyright:** © 2011 Lewis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was funded by the University of Pennsylvania University Research Foundation and the Nassau Research Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** in<sup>3</sup>, cubic inches

\* E-mail: lewisjas@stanford.edu



**Figure 2. Skull illustrated in Samuel George Morton's *Crania Americana* [8].** The 78 lithographs in *Crania Americana* set new standards for accuracy in anatomical illustration, as they were drawn carefully to scale and rechecked for accuracy multiple times, chiefly by John Collins [12,13]. Indeed, the quality of the illustrations in this volume exceeds that of many modern publications. The remaining specimens in the Morton Collection are currently curated at the University of Pennsylvania Museum of Archaeology and Anthropology. doi:10.1371/journal.pbio.1001071.g002

Morton again drew wide attention when Stephen Jay Gould used his skull research as a case study, first in a 1978 *Science* paper [1] and then in his 1981 book *The Mismeasure of Man* [5]. But this was no benevolent rescue from the backfiles of history. Gould reexamined Morton's data on cranial capacity variation in modern human populations and concluded that Morton had selectively reported data (see Box 1), manipulated sample compositions (see Box 2), made analytical errors, and mismeasured skulls in order to support his a priori views on intelligence differences between human groups. When properly analyzed, Gould argued, Morton's measurements show only trivial differences between populations. Gould

used Morton as a case study to argue that "unconscious or dimly perceived finagling, doctoring, and massaging are rampant, endemic, and unavoidable in a profession that awards status and power for clean and unambiguous discovery" [1]. Gould's analysis of Morton is widely read, frequently cited, and still commonly assigned in university courses [11–13]. Morton has become a canonical example of scientific misconduct and an oft-told cautionary tale of how human variation is inevitably mismeasured.

### The Current Study

The importance of the Morton case led us to reexamine the fundamental underlying

question: did Morton allow his a priori views on human variation to impact the data and analyses he published, as Gould argues? This hypothesis had remained essentially untested for 30 years. While some had raised questions about Gould's characterization of Morton [11–13], only one short publication by Michael [14] considered the primary data (Text S1). Unfortunately, the Michael study has multiple significant flaws rendering it uninformative (Text S1). It is rarely cited and, as noted by Kitcher, "virtually nobody has reacted to Michael's article by seeing it as a refutation of Gould" [15].

To test Gould's claim that Morton fudged his data, we relocated and remeasured almost half of the skulls that Morton had originally measured (Text S2, Dataset S1, Dataset S2). Gould did not measure nor personally examine the skulls in the Morton Collection—his argument was based on analyzing Morton's measurements. We also reanalyzed Morton's data and reexamined Gould's evaluation, drawing in part on the Stephen Jay Gould Archive recently made available. Our full analysis, along with all raw data, is given in the Supporting Information section (Text S1, Text S2, Dataset S1, Dataset S2, Dataset S3).

In reevaluating Morton and Gould, we do not dispute that racist views were unfortunately common in 19<sup>th</sup>-century science [6] or that bias has inappropriately influenced research in some cases [16]. Furthermore, studies have demonstrated that modern human variation is generally continuous, rather than discrete or "racial," and that most variation in modern humans is within, rather than between, populations [11,17]. In particular, cranial capacity variation in human populations appears to be largely a function of climate, so, for example, the full range of average capacities is seen in Native American groups, as they historically occupied the full range of latitudes [18]. It is thus with substantial reluctance that we use various racial labels, but it is impossible to discuss Morton and Gould's work without using the terms they employed.

### Remeasuring Morton's Skulls

Morton initially measured cranial capacity by filling skulls with seed, but he grew dissatisfied with the accuracy of this method and switched to using lead shot, which yielded more repeatable capacity values [8,10]. In Morton's initial seed-based 1839 study, "Caucasians" had the largest average cranial capacity (87 in<sup>3</sup>) followed by "Mongolians [Asians]"



**Box 1. Did Morton selectively report his data?** Morton divided his skull collection into broad racial groups, such as Native Americans, Caucasians, and so forth, but then also identified specific populations within those broad groups. So Morton's "Indian" (Native American) sample was composed of approximately 28 subsamples from more specific populations, such as Peruvians, Iroquois, "Eskimaux," and so forth [8]. One of Gould's best-known charges against Morton is that Morton was biased in his reporting of the cranial capacity averages for these subsamples: "It is intriguing that Morton often reported Caucasian means by subsamples, which permitted him to assert the superiority of Teutons and Anglo-Saxons. But he never broke down the Indian mean.... Thus, the fact that some Indian subsamples (Iroquois at 91.5 in<sup>3</sup>, N=4) exceeded the mean for Americans of Anglo-Saxon stock remained hidden in his raw data. (Morton never calculated the Indian subsample means at all; I have recovered them from his data.)" [1].

But Gould's claim, which has been repeated numerous times, is false. Morton routinely reported "Indian" subsample means, doing so at least 12 times in *Crania Americana*, the publication referenced by Gould. The subsample means reported by Morton include that of the Iroquois, which Morton noted was "within two inches of the Caucasian mean," as well as that of the "Eskimaux," which Morton noted was "a near approach to the Caucasian mean" [8]. In fact, Morton reported Native American population means more often than he did for other groups, and the means he reported are representative of his Native American sample as a whole.

(83 in<sup>3</sup>), "Malays [Island Southeast Asia]" (81 in<sup>3</sup>), "[Native] Americans" (80 in<sup>3</sup>), and "Ethiopians [Africans]" (78 in<sup>3</sup>) [8]. His final shot-based tally in 1849 again had "Caucasians" with the largest mean capacity (92 in<sup>3</sup>) followed by "Malays" (85 in<sup>3</sup>), the "Negro Group" (83 in<sup>3</sup>), and

the "[Native] American Group" (79 in<sup>3</sup>) [10].

Gould famously suggested that Morton's measurements may have been subject to bias: "Plausible scenarios are easy to construct. Morton, measuring by seed, picks up a threateningly large black skull,

**Box 2. Did Morton manipulate his samples?** Gould states that "as a favorite tool for adjustment, Morton chose to include or delete large subsamples in order to match grand means with a priori expectations" [1]. This criticism stems from the fact that each of Morton's broader racial samples (e.g., "Indian") were composed of multiple population subsamples, typically with differing mean cranial capacities. Thus it is possible to alter the overall "race" means by manipulating their constituent subsamples, and Gould charges that Morton did just that in order to obtain the results he expected.

For example, Gould compares the cranial capacities in Morton's 1839 and 1849 publications and finds that "Morton's Indian mean had plummeted to 79 in<sup>3</sup>.... But, again, this low value only records an increasing inequality of sub-sample size. Small-headed (and small-statured) Peruvians had formed 23 percent of the 1839 sample; they now made up nearly half the total sample" [1]. However, the "Indian" mean was 79.6 in<sup>3</sup> in Morton 1839 and 79.3 in<sup>3</sup> in Morton 1849, so the "plummet" Gould refers to was all of 0.3 in<sup>3</sup>. More importantly, Morton in 1849 [10] explicitly calculated his overall "Indian" average by taking the mean of three subgroups: Peruvians, Mexicans, and "Barbarous Tribes"—this is readily apparent in Morton's table reprinted in Gould [1]. As such, the percentage of the overall "Indian" sample composed of Peruvians is irrelevant to the overall mean, as it is only the Peruvian average which impacts the overall value. The Peruvian average changed by less than 1 in<sup>3</sup> from Morton 1839 ( $n=33$ ) to Morton 1849 ( $n=155$ ).

Clearly, Morton was not manipulating samples to depress the "Indian" mean, and the change was trivial in any case (0.3 in<sup>3</sup>). In fact, the more likely candidate for manipulating sample composition is Gould himself in this instance. In recalculating Morton's Native American mean, Gould [1] reports erroneously high values for the Seminole-Muskogee and Iroquois due to mistakes in defining those samples and omits the Eastern Lenapé group entirely, all of which serve to increase the Native American mean and reduce the differences between groups.

fills it lightly and gives it a few desultory shakes. Next, he takes a distressingly small Caucasian skull, shakes hard, and pushes mightily at the foramen magnum with his thumb. It is easily done, without conscious motivation; expectation is a powerful guide to action" [5]. While Gould offers this as only a "plausible scenario," and did not remeasure any crania, subsequent authors have generally (and incorrectly) cited Gould as demonstrating that Morton physically mismeasured crania (e.g., [15]).

We remeasured 308 of the 670 skulls (46%) whose capacity was published by Morton (Text S2, Dataset S1, Dataset S2). Linear and quantile-quantile regression identified Morton's measurements of 7 skulls (2%) as differing significantly from ours (Table 1), with a percentage difference in measurements of greater than 5.5% (Text S2). If Gould's hypothesis that Morton physically mismeasured some skulls due to racial bias were correct, we would expect the mismeasured crania to be non-randomly distributed by population. Specifically, we would expect Morton's overestimates to be concentrated on "white" crania, whereas his underestimates would be mostly "non-white" crania. We tested this using the binomial probability on population-quantile tables (Text S2) and found only one significant difference: Morton overestimated more Egyptian crania (3 of 13) than would be expected by chance. The overmeasured Egyptian skulls are specimens that Morton considered clearly "Negro," so his overestimation is obviously at odds with his predicted bias. Otherwise, Morton's errors were random with respect to population. Individually, Morton's three most overmeasured skulls are an Egyptian Copt that Morton considered "Negro" (+12%), a Seminole (+8%), and a "Native African Negro" (+7%). These results falsify the claim that Morton physically mismeasured crania based on his a priori biases.

## Seeds, Shot, and Bias

Gould's claim that Morton had mismeasured crania based on race derived from his comparison of Morton's seed-based and lead shot-based measurements, with different races experiencing different changes in their average cranial capacity between the two methods [1]. Gould reconstructs that in going from Morton's seed measurements to shot measurements the average capacity for different groups experienced different increases: 5.4 in<sup>3</sup> for Morton's black sample, 2.2 in<sup>3</sup> for his "Indian" sample, and just 1.8 in<sup>3</sup> for his Caucasian sample. Gould concludes that "surely something funny is

**Table 1.** Crania mismeasured by Morton with shot [10], using our measurements as the “gold standard.”

Specimen #	Population	Cranial Capacity (in <sup>3</sup> )		Difference	Measure Error
		Current	Morton		
761	Egyptian Copt	76	85	+12%	0.5%
754	Seminole	82	89	+9%	0.2%
994	Native African	71	76	+7%	0.4%
1435	Aymara	70	66	−6%	0.3%
949	Arickaree	80	75	−6%	0.2%
1326	Aymara	83	75	−10%	0.5%
70	Chetimaches	84	75	−11%	0.5%

Our capacity measurements (“Current”) have been adjusted to account for the average difference (about 4%) produced by the difference in our method versus Morton’s shot method (see Text S2). “Difference” is Morton’s measurement relative to ours. Specimens with a percentage difference of greater than 5.5% (more than 2.5 standard deviations from the mean percentage difference) are clear outliers and we consider them to have been mismeasured by Morton. “Measure Error” is our measurement error based on three repeated measurements of each cranium’s capacity.

doi:10.1371/journal.pbio.1001071.t001

going on here....I strongly suspect a systematic bias for undermeasurement of black skulls [during the initial seed-based measurements]” [1]. This is the evidence Gould offers in support of his “plausible scenario” that Morton may have physically mismeasured crania.

Morton only reported individual seed-based measurements for “Indian” crania, as they were the focus of his 1839 *Crania Americana* volume. Gould derived the “seed to shot” changes in Morton’s other samples by making guesses about which skulls had been included, rendering those values highly questionable (Dataset S3). For “Indian” specimens, however, the seed and shot measurements of specific crania can be compared directly. Gould made those comparisons and reports that the average increase from seed to shot is 2.2 in<sup>3</sup> [1]. But the average, the only result reported by Gould, is deceptive. We found that the changes from seed to shot measurements of individual crania range from an increase of 12 in<sup>3</sup> to a decrease of 10 in<sup>3</sup>, with a standard deviation of 2.8 (Dataset S3). These increases and decreases do not appear to be patterned by group. For example, one Peruvian cranium increases in capacity by 12 in<sup>3</sup> (+18%), while another Peruvian cranium decreases in capacity by 5.5 in<sup>3</sup> (−7%). This casts significant doubt on the hypothesis that mismeasurements with seed were a function of Morton’s racial bias.

Rather than bias, the source of changes between Morton’s seed-based and shot-based cranial capacities is more likely that stated by Morton himself: mistakes in the seed measurements. The seed-based measurements reported in *Crania Americana* were done in part by an assistant whom Morton later found

had made errors. Morton, in describing his 1849 shot-based measurements, stated, “All the measurements in this Catalogue [1849], both of the facial angle and internal capacity, have been made with my own hands. I at one time employed a person to aid me in these elaborate and fatiguing details; but having detected some errors in his measurements, I have been at the pains to revise all that part of the series that had not been previously measured by myself. I can now, therefore, vouch for the accuracy of these multitudinous data...” [10].

### Reevaluating Gould’s Analysis

Gould also performed his own analysis of Morton’s cranial capacity data and came to the conclusion that “there are *no* differences to speak of among Morton’s races” ([1], *italics in original*). For Morton’s 1839 seed-based measurements, Gould claims that Morton’s Native American average capacity is artificially depressed by his inappropriate use of a straight mean (taking the average of each individual specimen in the entire sample) rather than a grouped mean (first taking the average of each Native American population subsample, then calculating the mean of those means), since the former is sensitive to differences in sample sizes between “large headed” populations and “small headed” populations. In fact, the grouped mean for Morton’s Native American dataset is 79.9 in<sup>3</sup>, almost identical to the straight mean of 80.2 in<sup>3</sup> (Dataset S3). So Morton’s use of a straight mean actually slightly increased his Native American average. Gould’s calculation of a higher Native American average (83.8 in<sup>3</sup>) is entirely a function of Gould omitting 34 crania (of 144) as coming from

populations with samples of  $n < 4$  and, even by that criterion, erroneously excluding 6 crania, all with small cranial capacities (Dataset S3).

Gould’s reanalysis of Morton’s 1849 shot-based data resulted in a Native American mean capacity of 86 in<sup>3</sup> rather than Morton’s original 79 in<sup>3</sup> [1]. Gould obtained his new average by again taking the group mean of Native American populations with four or more crania. But Gould also applied an additional restriction: he only included Native American crania that Morton had also previously measured with seed. This restriction is entirely arbitrary on Gould’s part, as Morton’s publications and analyses for his seed- and shot-based measurements are completely separate (1839 versus 1849), and Gould did not apply this restriction to the other groups he reanalyzed in Morton’s shot-based data. If this restriction is lifted, Gould’s Native American average would be reduced to about 83 in<sup>3</sup>, considerably below his reported 86 in<sup>3</sup> (Dataset S3).

Overall, Gould concludes that his reanalysis of Morton’s shot-based data produces the “remarkable” result that there are no notable differences in mean cranial capacity between Morton’s groups, with Caucasians firmly mid-pack at 85 in<sup>3</sup> and the overall range being 83 to 86 in<sup>3</sup> [1]. However, Gould’s Caucasian figure was in error and should really be 87 in<sup>3</sup> rather than 85 in<sup>3</sup> [5]. And even accepting Gould’s inflated mean for Native Americans of 86 in<sup>3</sup>, the overall rank order of Gould’s results (whites/Native Americans/“Mongolians” and “Malays”/blacks) is then actually closer to Morton’s presumed a priori bias than were Morton’s own results (whites/“Malays”/blacks/“Mongolians”/Native Americans).

**Box 3. Did Morton use sex to skew his results?** Gould faulted Morton for failing to divide his samples by sex when calculating cranial capacity averages, given that differences in mean stature between males and females typically produce attendant differences in mean cranial sizes [1]. Certainly, more accurate population averages would be obtained if each sample were composed of equal numbers of males and females. But the question at hand is whether Morton manipulated his data to fit his preconceptions. In this regard, it is essentially impossible for Morton to have exploited sexual differences in cranial capacity to alter population averages. The only way Morton could have done so is by including more females for the populations he considered “inferior” and more males for “superior” populations. However, Morton did not collect the skulls himself [1,6], and there is no evidence that he excluded any skulls from measurement based on sex. Indeed, Morton was largely blind to the sex of the skulls in his collection because of the low accuracy of determining sex from the skull, a low accuracy noted as well by Gould [1]. Furthermore, given that Morton’s procedure for estimating sex from skulls almost certainly depended heavily on size (as noted by Gould, and as even modern methods do), it is entirely unsurprising to find a notable difference in size between “males” and “females.” Gould faults Morton for failing to notice this difference that “stared him in the face,” but had Morton commented on it he could rightly have been criticized for circularity—assigning sex based on size guarantees that “males” will appear larger than “females.”

## Our Verdict

Our analysis of Gould’s claims reveals that most of Gould’s criticisms are poorly supported or falsified. It is doubtful that Morton equated cranial capacity and intelligence [6,13], calling into question his motivation for manipulating capacity averages. Morton did not consider the influence of sex or stature on cranial capacity, but it would have been impossible for him to use those parameters to bias

the averages he reported (see Box 3). The grouped mean of Morton’s Native American sample is almost identical to the straight mean, rendering irrelevant Morton’s decision to use the latter. The changes in average cranial capacity from Morton’s seed-based measurements to shot-based measurements cannot be reconstructed with any certainty, incorporate erroneous seed measurements made by Morton’s assistant, yielded a broad range of changes ( $-10$  to  $+12$  in<sup>3</sup>) hidden

**Box 4. Did Morton ignore his mistakes?** Gould [1] found that in the final table of Morton’s main work, *Crania Americana*, Morton had erroneously reported the Native American mean cranial capacity as 82.4 in<sup>3</sup> rather than the true value of 80.2 in<sup>3</sup>. As Gould describes, “this elementary error permitted Morton to retain the conventional scale of being with whites on top, Indians in the middle, and blacks on the bottom” [1]. Gould argued that the error persisted because its “demotion” of blacks “provided so much satisfaction that Morton never thought of checking himself” [1]. However, the correct value is given on the page in *Crania Americana* preceding the table in question, suggesting the error in the table was typographical. Furthermore, historical evidence indicates that Morton did check himself and attempt to correct the error.

Michael [14] describes a copy of *Crania Americana* inscribed by Morton with the erroneous “82” value for “Indians” corrected in the same pen to read “80.” A different Morton-inscribed copy of *Crania Americana* reprinted by Bernasconi [21] has the same correction. We found that Gould’s personal copy of a first edition *Crania Americana*, while lacking an inscription from Morton, also has the identical correction in ink clearly of considerable antiquity (Gould Archive, Stanford University). In addition, Stanton [6] reproduces the same table with the correct value of 80 set in type. This suggests that a systematic effort to correct this error was made around the time of publication, casting doubt on Gould’s claim that Morton “never thought of checking himself.”

Finally, this error did not “demote” blacks: the rank ordering of groups by average cranial capacity remains “White/Indian/Black” whether “Indians” are 80 in<sup>3</sup> or 82 in<sup>3</sup>. As such, the error does not alter the “scale of being” whatsoever, contra Gould, falsifying the alleged motivation for Morton’s error.

by Gould’s mean, and are confounded by the shifts in sample composition (circa 50%) between the two rounds of measurement. Morton did not manipulate his samples to influence the average cranial capacities, at least not in a detectable manner. Morton did report subsample means for non-Caucasian groups (see Box 1). Of the approximately seven minor errors in Morton’s work identified by Gould [1], only two appear to be actual errors, and their overall impact confounds rather than supports Morton’s presumed a priori rankings.

Of the substantive criticisms Gould [1] made of Morton’s work, only two are supported here. First, Morton indeed believed in the concept of race and assigned a plethora of different attributes to various groups, often in highly racist fashion. This, however, is readily apparent to anyone reading the opening pages of Morton’s *Crania Americana*. Second, the summary table of Morton’s final 1849 catalog [10] has multiple errors (Dataset S3). However, had Morton not made those errors his results would have more closely matched his presumed a priori bias (and see Box 4). Ironically, Gould’s own analysis of Morton is likely the stronger example of a bias influencing results [11].

It should be noted that, were Gould still alive, we expect he would have mounted a defense of his analysis of Morton. We are saddened that his passing precludes such an exchange. While we differ with Gould in regards to his analysis of Morton, we find other things to admire in Gould’s body of work [19–20], particularly his staunch opposition to racism [5]. We trust that Gould, having reevaluated the work of Morton long after Morton’s passing, would find our reevaluation of “Gould on Morton” an appropriate exercise, even if he would likely have differed with our conclusions.

## Biased Scientists Are Inevitable, Biased Results Are Not

Samuel George Morton, in the hands of Stephen Jay Gould, has served for 30 years as a textbook example of scientific misconduct [12]. The Morton case was used by Gould as the main support for his contention that “unconscious or dimly perceived finagling is probably endemic in science, since scientists are human beings rooted in cultural contexts, not automatons directed toward external truth” [1]. This view has since achieved substantial popularity in “science studies” [2–4]. But our results falsify Gould’s hypothesis that Morton manipulated his

data to conform with his a priori views. The data on cranial capacity gathered by Morton are generally reliable, and he reported them fully. Overall, we find that Morton's initial reputation as the objectivist of his era was well-deserved.

That Morton's data are reliable despite his clear bias weakens the argument of Gould and others that biased results are endemic in science. Gould was certainly correct to note that scientists are human beings and, as such, are inevitably biased, a point frequently made in "science studies." But the power of the scientific approach is that a properly designed and executed methodology can largely shield the outcome from the influence of the investigator's bias. Science does not rely on investigators being unbiased "automatons." Instead, it relies on methods that limit the ability of the investigator's admittedly inevitable biases to skew the

results. Morton's methods were sound, and our analysis shows that they prevented Morton's biases from significantly impacting his results. The Morton case, rather than illustrating the ubiquity of bias, instead shows the ability of science to escape the bounds and blinders of cultural contexts.

## Supporting Information

**Text S1** Additional historical background. (DOC)

**Text S2** Materials and methods. (PDF)

**Dataset S1** Morton's raw cranial capacity data from his three major publications, *Crania Americana* [8], *Crania Aegyptiaca* [9], and *Catalogue of Skulls* [10], in Microsoft Excel format.

(XLS)

**Dataset S2** The raw cranial capacity data of the present study, in Microsoft Excel format.

(XLS)

**Dataset S3** The analytical spreadsheets showing the calculations described, in Microsoft Excel format.

(XLS)

## Acknowledgments

Thanks go to Alicia Bennette Harrison and Ivy Wilkinson-Ryan for their involvement in early aspects of this project, to Ian Robertson and James Holland Jones for statistical advice, to the staff of the Special Collections Department of Stanford University Library for assistance with the Gould Archives, and to the reviewers and editors for helpful suggestions.

## References

- Gould SJ (1978) Morton's ranking of races by cranial capacity: unconscious manipulation of data may be a scientific norm. *Science* 200: 503–509.
- Haraway D (1989) *Primate visions: gender, race, and nature in the world of modern science*. New York: Routledge.
- Kincaid H, Dupré J, Wylie A, eds (2007) *Value-free science? Ideals and illusions*. Cambridge: Oxford University Press.
- Marks J (2009) *Why I am not a scientist: anthropology and modern knowledge*. Berkeley: University of California Press.
- Gould SJ (1981) *The mismeasure of man*. New York: W. W. Norton and Company.
- Stanton W (1960) *The leopard's spots*. Chicago: The University of Chicago Press.
- Wake DB (2002) On the scientific legacy of Stephen Jay Gould. *Evolution* 56: 2346.
- Morton SG (1839) *Crania Americana; or, a comparative view of the skulls of various aboriginal nations of North and South America: to which is prefixed an essay on the varieties of the human species*. Philadelphia: J. Dobson.
- Morton SG (1844) *Crania Aegyptiaca; or, observations on Egyptian ethnography, derived from anatomy, history, and the monuments*. Philadelphia: John Pennington.
- Morton SG (1849) *Catalogue of skulls of man and the inferior animals, Third Edition*. Philadelphia: Merrihew and Thomson Printers.
- Brace CL (2005) "Race" is a four-letter word: the genesis of the concept. New York: Oxford University Press.
- Cook DC (2006) The old physical anthropology and the New World: a look at the accomplishments of an antiquated paradigm. In: Buikstra JE, Beck LA, eds. *Bioarchaeology: the contextual analysis of human remains*. Amsterdam: Elsevier. pp 27–71.
- Buikstra JE (2009) Introduction to the 2009 Reprint Edition of *Crania Americana*. Davenport, Iowa: Gustav's Library. pp i–xxxvi.
- Michael JS (1988) A new look at Morton's craniological research. *Curr Anthropol* 29: 349–354.
- Kitcher P (2004) Evolutionary theory and the social uses of biology. *Biol Philos* 19: 1–15.
- Judson HF (2004) *The great betrayal: fraud in science*. Orlando: Harcourt.
- Cartmill M (1998) The status of the race concept in physical anthropology. *Am Anthropol* 100: 651–660.
- Beals KL, Smith CL, Dodd SM (1984) Brain size, cranial morphology, climate, and time machines. *Curr Anthropol* 25: 301–330.
- Gould SJ (1977) *Ontogeny and phylogeny*. Cambridge: The Belknap Press of Harvard University Press.
- Gould SJ (1992) Impeaching a self-appointed judge. *Sci Am* 267: 118–121.
- Bernasconi R, ed (2002) *American theories of polygenesis*. Bristol: Thoemmes Press.
- Nott JC, Glidden GR (1855) *Types of mankind*. Philadelphia: Lippincott, Grambo and Co.

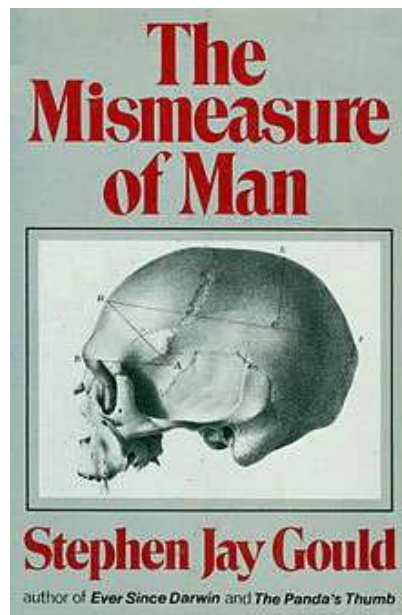


# Living Anthropologically

Anthropology – Understanding – Possibility

[Home](#) » [Blog Posts](#) » [Scathing Anthropology](#) » **Mismeasuring Gould**

## Mismeasuring Gould



## The Mismeasure of Science

In the 2011 *Plos Biology* “The Mismeasure of Science: Stephen Jay Gould versus Samuel George Morton on Skulls and Bias,” six anthropologists reassessed Gould’s famous 1981 book.

demands awareness and treading carefully.

## Anthropology and Gould

It appears Gould was wrong about Samuel Morton and his skull measurements. In “The Mismeasure of Science,” Morton becomes the careful and objective scientist, Gould becomes the ideologue. On his widely-read blog, [Gould’s ‘Unconscious Manipulation of Data’](#), John Hawks notes “the authors wrote in an even tone and lay out the facts in a very straightforward way. As a reader, I can’t see how they managed to keep their cool.” Hawks goes further: “This stuff really ticks me off. I don’t think that Gould’s errors can be written off as ‘unconscious bias’. Reading back over his 1978 article, I cannot believe that *Science* published it.” Hawks recommends “everyone should read [The Mismeasure of Science]. . . . It is a very suitable article for assignment in classes.”

Toward the goal of everyone reading this study, Nicholas Wade helped out with an article in the *New York Times*, [Scientists Measure the Accuracy of a Racism Claim](#). Wade has always pounced on anything that might validate traditional race categorizations. Wade leaves the last line for Ralph L. Holloway, one of the study co-authors: “I just didn’t trust Gould. . . . I had the feeling that his ideological stance was supreme. . . . I just felt he was a charlatan.”

I agree anthropology’s alliance with Gould has been problematic. First, using Gould in introductory textbooks makes anthropology look dated and out-of-touch with contemporary biology. Second, Gould consistently misinterpreted and misrepresented multiregionalism, which stymied an anthropological alternative to the Mitochondrial Eve hypothesis (see section [More mothers than Mitochondrial Eve](#)). However, this thorough trashing of Gould as an anti-science charlatan misreads the historical and contemporary context.

## **We read Gould because he was anti-racist, not anti-science**

I first read parts of the *Mismeasure of Man* in an undergraduate class on the history of imperialism. The point of reading Gould was not to impugn scientific inquiry, but to show the problems of harnessing scientific inquiry to racism and colonialism. I cannot of course do a systematic survey of how professors use Gould today, but I suspect Gould is still used more to illustrate how science must be careful to examine its premises and role in society, and less as a diatribe about how science is inevitably ideologically tainted.

There is no reason to doubt the study's claim that Morton was a quite objective and careful analyst. However, in going after Morton, Gould was tackling a kind of exemplary character for an entire era: Morton was measuring skulls and reporting results by a racial hierarchy, providing justification for colonialist and racist practices. As the study authors report, "we do not dispute that racist views were unfortunately common in 19th-century science."

But the problems extend beyond simply measuring and analyzing skulls—there is also the issue of how Morton acquired an enormous skull collection. Morton did not himself collect skulls, but he did promote their collection. As an official brochure on the [Morton Cranial Collection](#) notes, "the endeavor was so significant that army surgeons stationed in remote areas of the world would take great risks to obtain crania for Morton, and they were not above robbing graves in order to do so!" (It is interesting to note how often people use the passive voice to describe Morton's skulls, as in "were collected" or "were obtained.")

---

In the context of the time when Gould was writing *The Mismeasure of Man*, the scientific enterprise was indeed under severe critique for both participating in colonial racist practices and supporting those practices with analytical justifications. Gould's writings offered hope that scientific inquiry could correct past mistakes and excesses. At a time when science was under direct attack, Gould helped to save, rescue, and shield science as a potentially still-noble enterprise. (Gould was writing during a time when revisionist interpretations thoroughly questioned the traditional progressivist stance. This paralleled factors influencing Jared Diamond's [Worst Mistake in the History of the Human Race](#).)

## Debunking Gould in a climate of race resurgence

The contemporary climate is quite different. Today, science is not under the same kinds of attack. Instead, we are experiencing race resurgence and an assault on anti-racist views. The authors of this study seem unaware of the present climate and how their study plays directly into it.

The authors write as if race were already thoroughly debunked: "Studies have demonstrated that modern human variation is generally continuous, rather than discrete or 'racial,' and that



most variation in modern humans is within, rather than between, populations.” Here, they sound very much like Lewontin, and their sources for this statement are dated: they cite Loring Brace as well as Matt Cartmill, [The Status of the Race Concept in Physical Anthropology](#) (1998). Both are fine sources, but neither address the works of race resurgence, such as “Lewontin’s Fallacy” and the *New York Times* Op-Ed on “A Family Tree in Every Gene” (see blog-section on [Race Revival](#)). The authors use Brace’s 2005 book “[Race](#)” [Is a Four-Letter Word](#) several times, as an example of how Gould is still used, as someone who questioned Gould, and as a source for debunking the race idea. If Brace already accomplished all this, I would much use his book in class than this study.

If the authors are sincere that they “find other things to admire in Gould’s body of work, particularly his staunch opposition to racism,” then they need awareness of what is going on with regard to race resurgence and to help us with more contemporary references combatting that resurgence. A good place to start could be the articles in [Race Reconciled](#) in the 2009 *American Journal of Physical Anthropology* (for a summary, see “[Race Reconciled](#)” [re-debunks race](#)).

The take-away headline from this study is “Gould was wrong, Morton was right.” In a climate of race resurgence, this is treacherous territory. Nicholas Wade, of course, makes no mention of any race debunking or how the authors share Gould’s opposition to racism. Instead, Wade inserts this troubling line: “When [Gould’s errors] are corrected, the differences between the racial categories recognized by Morton are as he assigned them.” Without bothering to question Morton’s racial categories, Wade appears to assert that in fact Morton’s cranial measurements by race are validated.

## On a True Mismeasure of Science

The authors try to frame the larger framework of their study in terms of its impact on science studies. They claim the Morton case “shows the ability of science to escape the bounds and blinders of cultural contexts.” However, this is at best a minor sparring point for a dispute internal to academia. I say “at best,” because it is not even a sparring point for people doing a true history of science, for whom the question of “bias” has long been eclipsed by documenting particular practices and particular analytical techniques. (Kenan Malik has

some very nice reflections on this piece and what it means for science-society at his blog-post [The science of seeing what you want to see.](#))

Regardless of whether science can “escape the bounds and blinders of cultural contexts,” the results of science must still be presented within a particular historical and political context. The authors of this study seem blissfully unaware and naïve about that contemporary context. Nicholas Wade’s quick coverage should be a wake-up call. In a context of race resurgence, and without providing contemporary support for anti-racist scholarship, this study will soon ricochet through the conservative blogosphere—and many undergraduate classrooms—as “Gould was wrong, Morton was right; cranial capacity and race is like Morton said it was.”

---

**Update 2013:** I received an e-mail alert from the undergraduate who originally measured the Morton skulls and who has posted [a 4-part personal commentary on Stephen Jay Gould and Samuel George Morton](#). Looking through that material, and the kinds of commentaries it prompts, and then back at this post leads me to re-state the following points:

1. In our contemporary climate—see [White-Race Problems](#) and [Social Construction of Race = Conservative Goldmine](#)—the fundamental message that “Gould was wrong, Morton was right” can only be fodder for racism and the race resurgence.
2. Related to that point, I have yet to see anything creative about the Morton skull collection, as in “measuring the Morton skulls tells us these new facts about human variation.” The re-measurements seem only to be about Gould on Morton, rather than on what we can say about human variation.
3. Again, most people read Gould because he was anti-racist, not anti-science, and at the time Gould’s analysis served to save the scientific enterprise from a variety of stronger criticisms.

---

To cite: Antrosio, Jason. 2011. “Mismeasuring Gould in “The Mismeasure of Science.” *Living Anthropologically* website, <https://www.livinganthropologically.com/mismeasuring-gould/>. First posted 14 June 2011. Revised 7 September 2017.



□ Scathing Anthropology  
 □ ambushing anthropology, anthropology branding, evolution, genetics, Nicholas Wade, politics, race, racism

---

**Living Anthropologically means documenting history, interconnection, and power during a time of global transformation.** We need to care for others as we attempt to build a world together. This blog is a personal project of Jason Antrosio, author of [Fast, Easy, and In Cash: Artisan Hardship and Hope in the Global Economy](#). For updates, follow on [Twitter](#), watch on [YouTube](#), or [subscribe](#) to e-mail list.

# WHO COUNTS?

## The Politics of Census-Taking in Contemporary America

Margo J. Anderson ■ Stephen E. Fienberg

## Chapter 2

---

# The History of the U.S. Census and the Undercount

In the summer of 1787, slightly more than a decade after the thirteen colonies had declared their independence from Great Britain, several dozen men met in Philadelphia to try to improve the existing American national government structure. War had ended in 1783, and the infant nation had returned to peace. Yet severe political and economic problems plagued the country. The Articles of Confederation, finally ratified in 1781 as the framework for the national government, had not been functioning well for a number of years. By the late 1780s, the states were willing to send delegates to discuss amendments. There were many different proposals about what to do, and although these men had the common experience of the Revolution to unite them, they also had sectional, political, religious, and local interests to divide them.

The men who gathered in Philadelphia thus faced a delicate and complex set of political questions. On the one hand, they had to empower a national government to deal with recurring problems arising from the unique historical development of the United States. The governments of the individual colonies had had few connections with one another prior to independence; they did not easily merge into a United States and, in fact, often looked jealously upon one another. One set of issues revolved around replacing the functions the British had served during the colonial era: national defense, diplomacy, trade policy, taxation, the opening of new land, and the creation of new colonies. On the other hand, the national government had to respect the autonomy of the individual states, recognize and mediate differences among the states and their citizens, and develop mechanisms to

apportion power and tax burdens among the constituent elements of the United States. Revising the national government or devising a new one would be no easy matter. Several prior efforts had failed. Contemporaries might well have expected the men to disband with little accomplished at the end of the summer. We know now that that did not happen, that their deliberations resulted in a fundamentally new and breathtakingly radical governmental structure, which we have lived with ever since. The Constitution created three branches of government—executive, legislative, and judicial. (The Articles of Confederation had mandated a one-house congress and no executive branch.) The legislature was bicameral, with the lower house apportioned among the states according to population and elected by the people, the upper house apportioned among the states and elected by the state legislatures. The executive branch was headed by a president elected by the people through the mechanism of the electoral college. The new system increased the power of the national government considerably, yet it was also full of checks and balances to guarantee the powers of the states and citizens and to protect against the kind of tyranny the Crown had exercised.

The complex and relatively explicit structures, the detailed enumeration and delegation of powers, the checks and balances—all were mechanisms to foster both national unification (“a more perfect union”) and the rights of the individual states. So also was the institution of a decennial census to measure the relative strength of the various elements of the population and hence to periodically readjust the relative power of the states and local areas in the national government. Article I, Section 2, of the Constitution created the House of Representatives and defined its membership and capacities. Paragraph 3 of Section 2 specified that the representatives and direct taxes were to be “apportioned among the several States which may be included within this Union, according to their respective numbers.” This simple language provided the solution to one of the fundamental political controversies of the revolutionary era: namely, how to allocate representation fairly in legislative assemblies. Before the Revolution, the colonists had protested their lack of representation in the British Parliament. The newly united thirteen colonies had also struggled over the problem of the equitable distribution of the burdens and resources of the national government among large and small states. Critics had charged that the Articles of Confederation were an unsatisfactory grounding for a national government precisely because states had voted as units, regardless of their disparate wealth or populations. The 1787 Constitu-

tion allocated representation in the House according to population and in the Senate by state. Since direct taxes were also to be allocated according to population, the large states would gain House representation but incur a higher potential tax burden to the federal government. The strength of each state in the electoral college was to be determined by summing the state's Senate and House members. Finally, the census was to be taken every ten years, because the framers were well aware that populations—and especially the American population—grew and shifted over time.

The logic of the census system flowed from the experience and conceptions of the framers. The Founding Fathers debated other methods of allocating political power and tax responsibility during the Philadelphia convention. They discussed apportionments based on land assessments, other measures of wealth, and population. They agreed that theoretically, political power should be allocated on the basis of population and that tax responsibility should be based on wealth. Everyone agreed, however, that population was much easier to measure than wealth and that the wealth of states was highly correlated with population. Thus, population would be the apportionment measure. The new census and apportionment mechanisms of the federal Constitution were thus a crucial piece of the “Great Compromise” between the large and small states that made a national government possible. The periodic changes that had to take place to account for population growth and change were assigned to an automatic decennial routine, separated from the cycles of more frequent elections for House, Senate, and the presidency—which were also designed to shift power among the constituent elements of the population.

But there was one fly in the ointment that would come back to haunt the framers in future years and that is of particular relevance to us here. That was the question of defining exactly who was part of the “population” deserving the right to political participation in the society and owing responsibility to pay taxes to the state. The fundamental thrust of the discussion in the Constitutional Convention was to use the most expansive rule possible, including, for example, women, children, and the poor in the count, though they neither voted nor were necessarily responsible as individuals for taxes. The rub came when the framers considered how to treat slaves and Indians. Should the southern states, for example, be granted political representation for their slaves? At the time, southerners considered slaves property for purposes of tax assessments but did not count slaves when they apportioned their state legislatures. The double rule of using the same

measure for representation and taxation broke down logically for slaves. Furthermore, they needed to consider whether Indians, who were generally considered outside the purview of the American polity and as members of foreign states, should be counted for representation and taxation.

The solution to these dilemmas was to hedge the universal rule of counting the population for apportionment with two provisos. The census clause in Article I, Section 2, paragraph 3, continued: the “respective Numbers [of the population] . . . shall be determined by adding to the whole Number of free Persons, including those bound to Service for a term of Years, and excluding Indians not taxed, three-fifths of all other Persons.” The three-fifths compromise required the census to count slaves separately so they could be considered, for apportionment, as three-fifths of a free person. The second proviso eliminated “Indians not taxed” from the census altogether. Only people who came to be called “civilized Indians” were to be included in the decennial census count and, hence, in the apportionment totals. The important point to be made about these constitutional provisions was that they defined the population to be included or excluded, counted as whole people or as three-fifths of a person, according to the person’s civil status. To avoid ambiguity, the Constitution even clarified that unfree indentured servants, that is, “those bound to Service for a term of Years,” were to be included in the category “free Persons.” Nowhere did the Constitution mention a racial classification, and in fact, the framers used the ambiguous “other persons” to define “slaves.” But very quickly, as we shall see, in popular discourse and in the variable names listed on the census form, the racial attributes “White,” “Black,” “colored,” and “Indian” replaced the civil statuses of free, slave, and “Indians not taxed.”<sup>1</sup>

## Demographic History

Today, representative democracy is a generally accepted principle, and nations around the world take periodic censuses. A census seems a fairly obvious tool to use to apportion political power among a set of constituencies. Yet the United States was the first nation in the world to institute a regular population count to apportion political power. The principles that political power was a function of population and that population could be measured were truly innovative in the eighteenth century and proved to be lasting only as they were implemented in the nineteenth and twentieth centuries.



Moreover, the census might well have been a rather minor constitutional innovation were it not for the extraordinary demographic character of the American population. The United States has had one of the most heterogeneous and demographically dynamic populations in the history of the world. In the past three centuries, the land that became the United States has seen rapid population growth, major migrations, and sharp demographic transitions—all in the context of a racially and ethnically heterogeneous population. In 1700, about 250,000 people lived in the colonies; by the time of the Revolution the population exceeded 2 million. The first census counted 3.9 million people, and the current population is more than a quarter billion (see table 2.1). The current land area of the country is four times the size of the nation in 1790; the population is almost sixty times larger. In 1850, the country was 85 percent rural; now it is 74 percent urban. From 1700 until the Civil War, the American population grew at the rate of 30 to 35 percent a decade. From 1860 to 1910, it grew about 24 percent a decade; since then it has grown about 13 percent a decade. The median age of the population has increased greatly. In the early nineteenth century, it was around sixteen, reflecting the high birthrates and shorter life expectancy of the times. Now, it is around thirty-four. In the early nineteenth century, a child could expect to live to about the age of forty; now, a child will live for about seventy-five years.

The population of the United States was and remains racially and ethnically diverse. At the first census in 1790, about a fifth of the population was African or African American and primarily slaves. The larger majority was free and “White”—primarily from the British Isles. Although slavery ended during the tumultuous Civil War of the mid-nineteenth century, racial distinctions continue to be major social markers in American society. Currently, about 12 percent of the population is Black.

Between 1820 and 1980, fifty million immigrants came to the United States in search of jobs and a new life; during the years of the major European migrations of the late nineteenth and early twentieth centuries, 13 to 15 percent of the population was foreign born; currently 9.6 percent are foreign born. Approximately 1.4 percent of the current United States population is estimated to be undocumented immigrants. Today, 11 percent of the population is of Hispanic origin. This is a diverse population made up of recent immigrants and Americans of Hispanic ancestry. Finally, the American population was and is regionally heterogeneous. For example, currently 74 percent of the

**Table 2.1** Resident Population of the United States

Year	Population
1790 (Aug. 2)	3,929,214
1800 (Aug. 4)	5,308,483
1810 (Aug. 6)	7,239,881
1820 (Aug. 7)	9,638,453
1830 (June 1)	12,866,020
1840 (June 1)	17,069,453
1850 (June 1)	23,191,876
1860 (June 1)	31,443,321
1870 (June 1) <sup>a</sup>	39,818,449
1880 (June 1)	50,155,783
1890 (June 1)	62,947,714
1900 (June 1)	75,994,575
1910 (April 15)	91,972,266
1920 (Jan. 1)	105,710,620
1930 (April 1)	122,775,046
1940 (April 1)	131,699,275
1950 (April 1)	151,325,798
1960 (April 1)	179,323,175
1970 (April 1)	203,302,031
1980 (April 1)	226,542,199
1990 (April 1) <sup>b</sup>	248,709,873

Source: U.S. Bureau of the Census 1996c.

<sup>a</sup>Revised to include adjustments for underenumeration in southern states.

<sup>b</sup>Unadjusted.

population is urban. However, in thirteen states, more than 80 percent of the population is urban; in another seven states, the population is more than 50 percent rural. Populations in seven states (all in the South) are more than one-fifth Black, and in twenty-three states the Black population is less than 5 percent. Overall, 9.6 percent of the population is foreign born; yet the populations of New York, New Jersey, Florida, California, and Hawaii are 10 to 15 percent foreign born. In another twenty-four states, the population is less than 3 percent foreign born.<sup>2</sup>

These dramatic patterns make the decennial census and its apportionments major social, political, and intellectual events. Americans

are used to substantial demographic changes each decade and look to the census for their evidence. Census-based apportionments are supposed to take some difficult questions about the distribution of political power and economic resources off the immediate legislative agenda and consign them to predetermined allocations. Yet because so much is at stake, the census also runs the risk of being politicized. The census numbers are subject to intense scrutiny and analysis as each decade Congress, state legislatures, and other governmental bodies redistribute legislative seats, tax revenue, and grants-in-aid on the basis of the population-based apportionment formulas. By definition, for every gainer in the apportionment game, there must be a loser; legislators can soften the pain for the losers by shifting relative, but not absolute, power or revenue. Nevertheless, ultimately relying on the census leads to a zero-sum reapportionment game, which, though necessary, is still politically delicate and sometimes painful. Americans have not always agreed on whether the demographic changes in the population are good or bad, implying "progress" and "manifest destiny" or "degradation" and "race suicide." Accordingly, the development of the census itself as a more elaborate and more scientific count each decade is intimately bound up with the political and social history of the nation.

### The Development of a Census Instrument, from 1790 to 1840

From a late-twentieth-century perspective, we know that a host of methodological issues are involved in taking a successful census, and another host of methodological issues are involved in devising a "fair" system of apportionment. From our perspective, the history of successful census taking turns on the achievement of an accurate count, given the technical capabilities of the government at the time, and thus the willingness of Americans to see the census as a legitimate and essential part of the federal governmental machinery.

The same men who wrote the Constitution discovered as much when they found themselves as legislators and officials in the new federal government. One of their first tasks was to write a bill to take the 1790 census. James Madison proposed a rather elaborate census, which included questions classifying the population by age and sex and a census of occupations, instead of the simple count of the free and slave population required by the Constitution. Congress rejected Madison's proposal and settled on a more modest scheme. The first

census law mandated the secretary of state to charge the U.S. marshals to count the population in six simple categories. For each household, the enumerator listed the name of household head; the number of free White males under sixteen; the number of free White males sixteen and over; the number of free White females; the number of free colored; and the number of slaves. Even this simple count took eighteen months to complete as the difficulties of counting the population in rural America emerged. The first census recorded 3.9 million people. President George Washington thought the count was too low; but it was credible enough to be used for apportionment, and so Congress took up the second phase of the decennial census process: reapportionment.

At this point, as Michel Balinski and H. Peyton Young (1982) have indicated, another set of issues emerged. The Constitution was silent on the method of apportionment. Soon, Treasury Secretary Alexander Hamilton and Secretary of State Thomas Jefferson, the two leaders of the emerging political factions in the new government, came to be identified with two different methods of apportionment. Both methods were mathematically valid, but they had different effects on the distribution of seats in the House. A long debate ensued before Congress passed Hamilton's bill. George Washington sought advice from both men and vetoed the bill (using the presidential veto for the first time). At this point, everyone recognized that a constitutional crisis lurked behind the arcane issues of the census. If the legislators were not able to "constitute" Congress, there would be no government. Some weeks later, Congress passed a second bill, one that favored Jefferson's method. George Washington signed it, and all parties put the issue behind them for the decade.

The lessons the framers took from their initial experience with the census and apportionment were several. First, the census, and hence the new constitutional process, "worked." Furthermore, the framers recognized, there were questions of fairness inherent in the census and apportionment mechanisms, which the Constitution did not address. Third, it was perhaps best not to tinker with the process. The methods of the 1790s were thus used for the next fifty years, with minor modifications. Neither Congress nor the president saw the need to do much to improve the process. Contracts for the few clerical jobs involved in overseeing the census in Washington were given to political supporters of the administration; in a period of increasingly partisan politics, the most exciting facet of the census process was the congressional debate over the printing contract award. Over time, though, both the

possibilities and the problems with the census began to capture the attention of politicians and the mathematical tinkerers of the early republic. After a few decades, one could calculate growth rates for the nation, the states, and local areas. Almanacs reprinted the latest figures. Congress could ponder the changing structure of the various states and factions in Congress. The data seeped into the political discourse of the nation. Their use, in turn, led to further curiosity about the character of the population. Each decade, Congress refined the age and sex breakdowns in the census and added a few new questions. By 1840, the six simple questions of 1790 had expanded to seventy or more. And, more ominously, the growing sectional split between North and South echoed through the decennial census and apportionment process. The free states grew faster than the slave states, and hence free-state representatives increasingly dominated the House. Votes in the electoral college similarly tilted toward the North. Southerners watched these trends with great fear and began to articulate positions that would allow them to opt out of a government dominated by hostile northerners or to devise mechanisms—such as the creation of new slave states—to restore their strength in the House and electoral college.

### Census Accuracy: Sectionalism and the 1840 Census

At the time of the 1840 census, these trends were juxtaposed as the first major controversy over census accuracy erupted around the reported differential rates of insanity for free Blacks and slaves. The 1840 census included questions asking how many members of the household were “insane and idiotic.” “White” and “colored” insane and idiotic were to be separated. The numbers were duly summed in the published census results and reported by states. Southern papers reported the data and noted that the rates of insanity for Blacks were much higher in the North than in the South. They concluded that freedom drove Blacks insane and argued that slavery was therefore the appropriate civil status for the Black population. Needless to say, antislavery leaders in the North objected strenuously and began to look at the data on which these assertions were based to find mistakes. Northerners initially accused Secretary of State John C. Calhoun of fabricating the data. Calhoun denied this but refused to authorize an independent inquiry into the issue. Edward Jarvis, a young Boston physician, pointed out that the data had to be in error, because the

census reported insane free blacks in some Massachusetts towns that had no Black population (Cohen 1982).

The controversy raged for several years. John Quincy Adams pressed for a congressional investigation. The data were never changed, but the debate had several effects on the future of statistics and the census. First, Congress recognized that the accuracy of the census could be challenged. Second, Congress recognized that the challenge could be congruent with other political controversies—in this case, the debate about the future of slavery. Thus, the very system designed to depoliticize the apportionment of power in the United States could itself become part of the controversy. Third, Congress decided that the solution to these problems was to reform the census—that is, to take a better census in order to avoid errors in the data. To take a better census, Congress had to turn to experts who could suggest improvements. The 1850 census would be a much more ambitious undertaking, guided by the advice of statisticians who could guarantee that the 1840 problems would not recur.

### Reforming the Census: From 1850 to 1890

In 1850, in the midst of the contentious debates about the future of slavery in the territories, Congress undertook a major overhaul of the census law and, in so doing, restored confidence in the census process. Before 1850, the census used the family as the unit of interest and reported few data on persons. The change to a focus on individuals in census taking was strongly influenced by the work Lemuel Shattuck—one of the American Statistical Association's founders, who had earlier conducted the Boston census of 1845—as well as that of Adolphe Quetelet, who helped to organize the 1846 Belgian census.

The United States instituted an individual-level census; created a large, but temporary, census office to tabulate the data; and embarked on a marked expansion of the publications from the census. The office, housed in the new Department of Interior, grew rapidly. In the mid-nineteenth century, the Census Office employed about 10 percent of the federal employees in Washington when the census processing was in full force—testimony to both the elaborate scale of the census and the relative lack of other federal government activity at the time. There was a good deal of grumbling in the press and in Congress about this rapid expansion, but because the patronage jobs the office dispensed proved to be advantageous to politicians, they were not anxious to

certain appointments. In the years that followed, patronage continued to be a central feature of census taking, but the size of the census staff never returned to the pre-1850 scale.

The American population rose from twenty-three million in 1850 to fifty million in 1880. Congress called upon the Census Office to collect and publish more data on the demographic, social, and economic character of the population. The 1850 census-reform law had created six schedules for the statistics of population, mortality, agriculture, manufacturing, and social statistics. The data derived from these schedules were tabulated and published as separate census volumes and summarized in a shorter abstract or a compendium. Surprisingly, perhaps, the Civil War itself did not disrupt either the 1860 or the 1870 censuses. The counts were taken, tabulated, and published during the war and Reconstruction, though we know now that the 1870 field enumeration in the South was poor. Nevertheless, at the time, the census again seemed to be the instrument that documented national growth and development, not sectionalism and fragmentation. The former Civil War general and Mugwump Republican Francis Amasa Walker replaced the Free Soil Whig Joseph C. G. Kennedy as census superintendent in 1870, and Walker became the key census innovator of the post-Civil War era.

Once the Washington Census Office existed, Congress, interested public officials, and the clerks responsible for the census recognized the weaknesses of the U.S. marshals and their assistants as the field staff for taking the census. Up to this point in time, the marshals had been in charge of all field operations, including mapping their districts and appointing sufficient assistants to actually do the canvass. Needless to say, the accuracy of the enumeration relied heavily on the abilities of the marshals—who otherwise worked for the attorney general's office, not the Interior Department, which oversaw the census. The census superintendent had no real bureaucratic control over the marshals. Sloppy work, late work, or incomplete work was hard to correct. There was no training, no means for teaching someone how to take the census except by the detailed instructions issued from Washington. One wonders how frequently the fine print was read. Furthermore, marshals were assigned according to the needs of the attorney general, not the needs of the census, and the marshals and everyone else complained of uneven work loads. In 1870, a marshal in New York supervised assistants who counted more than two million people. Another in South Florida supervised assistants who canvassed an



underpopulated frontier area with only sixty-five hundred people but with many smugglers.

The first efforts at changing the census field staff occurred during Reconstruction. They were initially a side effect of a congressional debate about the impact of the abolition of slavery on the census and apportionment. The Republican congressman James Garfield, of Ohio, began an investigation of census legislation to see what other changes in census law might be necessary. His committee worked in the summer of 1869 and eventually proposed a wide-ranging reform of the census machinery, including the creation of a field staff—supervisors and enumerators—appointed by and responsible to the census superintendent. Garfield's bill passed the House but failed in the Senate in early 1870 because of the field-enumeration provisions. Marshals were appointed by senators; Garfield's supervisors would be allocated by House districts and presumably appointed by House members. In a day when patronage meant so much in Washington, the patronage shift was enough to kill the bill, and the 1870 census was taken under the old system. Francis Walker ran the 1870 census as efficiently as he could from Washington, all the time pointing out the limitations of a system that used U.S. marshals as the field staff.

Walker and Garfield also recognized that they would have to do a bit more political homework to get a bill passed for the 1880 census. Walker expanded the 1870 clerical staff considerably, but he made sure he stayed within the limits of the budget for the 1870 census; he ran the office efficiently and even disbanded it three years into the decade to avoid charges of boondoggling. Late in the 1870s, he and Garfield came back with a revised census bill. Under the new bill, field-staff supervisors and enumerators reported to the census superintendent. However, the supervisors were also appointed by the president, with the advice and consent of the Senate. In 1879, the bill passed relatively easily. Whereas the field enumeration in 1870 had been run by about 65 marshals, in 1880 there were 150 supervisors. No enumeration district was to be larger than four thousand people. The time for the enumeration was shortened from six months to two weeks in cities and one month in rural areas. Walker hired Henry Gannett, from the Geological Survey, to be census geographer. Gannett systematically mapped the country for the census, set up a Geography Division in the Census Office, and began the collection of a complete set of accurate maps of minor civil divisions around the country. This alone was a major administrative and technical innovation, since the Geological Survey was still in its infancy and many

parts of the country west of the 100th meridian were not mapped at all. The superintendent was beginning to exert more precise control over the conduct of the census in the field, but these efforts still left much to be done in the future: Correspondence was still by and large written in longhand, with letterpress copies for office records. The Geography Division and what we would now call the Field Division (then the Division of Supervisors' Correspondence) were separate entities and remained so. As in the past, the major control the central office had over supervisors and enumerators was financial. Before enumerators could be paid, their returns had to be certified as complete and error free by the supervisor and the Washington office. Communication back to the field was still extremely slow. Only rarely did an official from Washington visit the field.

Walker was the first census superintendent to confront complaints of local officials about census undercounts. The most spectacular of the challenges involved New York and Philadelphia. In 1870, both cities demanded and got recounts; in both cases, the second count came in about 2 percent over the first. Indianapolis also complained in 1870 when the census did not show its population at fifty thousand. A second count was done there, too—but the city boundaries were expanded considerably to increase the population by 19 percent. The complaints continued in later years: in 1890, New York City (a Democratic city) again complained that its population had been undercounted and blamed the Republican administration in Washington and the local supervisor; 2 percent of the enumerators were assigned to reenumerate districts “where the original work had been improperly done” (Holt 1929, 29). These partisan challenges to census accuracy led to much political verbiage but did not affect the process of census taking or fundamentally undermine the instrument itself.

## The Census at the Turn of the Twentieth Century

The process of growth that had begun in 1870 led to a radically different census by 1900. At the same time, reapportionment, which had dominated public attention toward the census through much of the nineteenth century, receded as a controversial public issue. Before the Civil War, easterners had conceded the shift in political power to the growing areas of the West. After the Civil War, southerners conceded that their demographic decline would lead to a smaller role in Congress. Ironically, this decline was eased somewhat by the

elimination of the three-fifths rule. Southern states received a boon of additional representation when the four million slaves were freed. In the postbellum era, Congress “solved” the problem of differential population growth by increasing the size of the House at each census. In the 1860s, Congress had 243 members; by 1910, there were 435. During that period, no state lost a seat through reapportionment. As the science of statistical data collection grew, census officials concentrated on improving their capacity to collect, process (classify and tabulate), and publish data.

The 1900 census was something of a high point for the expanding census. Congress, about to make the agency permanent, was willing to recognize that taking the census required special expertise. The 1900 census law authorized the appointment of an assistant director, “who shall be an experienced practical statistician,” and five chief statisticians, “who shall be persons of known and tried experience in statistical work.” William Merriam, the 1900 census director, was a Republican political appointee in the older tradition, but the list of statisticians, including Walter Willcox, S. N. D. North, Joseph Hill, Wesley Mitchell, and Allyn Young, read like a Who’s Who of early-twentieth-century statistics and social science. They, among other 1900 census officials, all served as president of the American Statistical Association in later years. Many commentators at the time—surveying the growth of statistical activity throughout the federal government—even saw the Census Office on its way to becoming a true central statistical bureau, on the European model.

The Census Bureau became a permanent federal agency in 1902, located in the new Department of Commerce and Labor, but Congress did not see fit to centralize the federal statistical system. The new permanent Census Bureau, though now charged with collecting data on a wide variety of topics throughout the decade, fell upon difficult political times by 1910. Bureau leaders continued to press for centralization of federal statistical activities under the banner of an expanded Census Bureau. But statisticians in other federal agencies resisted the bureau’s goals of consolidating all federal statistical programs in one agency. Between 1900 and 1910, it became clear that officials of the Bureau of Labor Statistics, and statisticians in other cabinet offices, such as the Departments of Agriculture and Treasury, saw things differently. Their statistical programs had also grown in the last thirty years of the nineteenth century, and they were integrated into the larger administrative functions of their agencies. These officials saw

no reason to transfer their data collection and publishing activities to an agency in another cabinet department.<sup>3</sup>

When younger statisticians recognized that the Census Bureau was not going to become a true central statistical bureau, they saw it as a less attractive place to look for a career. Furthermore, the bureau did not continue to develop a job hierarchy. There were few mid-level jobs and fewer long-term possibilities for promotion. The overwhelming number of positions were clerical; positions contracted rapidly after the decennial census buildup. Innovation in statistics in this period generally began to shift from government agencies to the new research universities, with their laboratories and consistent research agendas. Finally, the arrival of the Wilsonian Democrats in 1913 prompted even more staff turnover in this traditionally Republican agency. Thus, the Census Bureau settled into a permanent routine existence as one of many bureaus within the Department of Commerce. (The Labor Department got separate cabinet status in 1912.) The decennial census, though still the major activity of the agency, competed with new statistical initiatives in other areas: manufacturing statistics, cotton statistics, vital statistics, special analytic studies, and contract work for congressional interests such as the Immigration Commission.

World War I created new data needs that reinforced these administrative trends. American entry into the world war severely challenged the federal government's ability to produce information for the war effort. The limitations of the uncoordinated prewar statistical system were quickly obvious to Congress and the administration. The nation needed accurate estimates of manpower, manufacturing capacity, agricultural strength, and naval and shipping capacity. Different agencies produced different numbers; eventually, President Woodrow Wilson consolidated the nation's efforts under Edwin Gay's Central Bureau of Planning and Statistics. Many former census officials, including a number of the chief statisticians from 1900 and 1910, returned to work on the statistical war effort. By the end of World War I, these efforts led to improvement of the situation in the Census Bureau and the development of something of a formal advisory and lobbying arm among professional statisticians. Longtime census officials proposed the creation an advisory committee composed of statistical experts—three each appointed by the American Statistical Association and the American Economic Association. Often composed of former census officials or university faculty who were strong supporters of the Census Bureau, this committee began to meet in 1919 and, in various forms including separate committees, has met several times each year

ever since. Advising the bureau on policy, questionnaire design, and other aspects of statistical innovation, the Advisory Committee has become a tangible institutional link between the Census Bureau and the growing professional statistical community.

These trends continued during the 1920s, as the Census Bureau became a key agency in Herbert Hoover's Commerce Department. The Advisory Committee became a kind of unpaid policy board and fit well with Hoover's theories of the way the private sector and voluntary organizations could foster the "associative state" and expand bureau programs without the heavy hand of government planning. In the 1920s, the bureau focused especially on business statistics, developing the Survey of Current Business and the Census of Distribution. There was little talk of creating a central statistical bureau. Nor did the organizational structure within the bureau change much. Census monographs were contracted out or subsidized by private organizations. Innovations in the field of statistics increasingly took place in university settings rather than within government, as had occurred in the nineteenth century.

The chief controversy surrounding the census during the 1920s concerned the failure of Congress to reapportion itself after the 1920 census. For the only time in the nation's history, Congress could not muster a majority to pass a reapportionment bill. The roots of the problem were several. First, Congress had decided in 1910 to stop the growth in the size of the House. Thus, the 1920 reapportionment would be a true zero-sum game: one state's gain would be another state's loss. Second, several mathematicians produced competing apportionment methodologies, which had the effect of producing different allocations for key states. Third, the demographic trends evident in the 1920 census results were not to the liking of the Republican majority in Congress. The census results showed major shifts in population to the cities, to the far West, and to places in the country populated by immigrants. The census, critics held, had to be wrong. Leaders proposed, and failed to pass, bill after bill until, by the late 1920s, key congressional leaders proposed an automatic mechanism to go into place after the 1930 census. This bill passed Congress with another provision, which removed the mandate—present in every apportionment bill since 1840—that required congressional districts to be compact, contiguous, and, as nearly as possible, equally sized within states. Congress, in other words, reapportioned power among the states but quietly acquiesced to rural malapportionment within states. States with growing urban populations would gain in the number of

seats, but those seats would not necessarily be allocated to urban districts. It was a practice that would lead to the reapportionment revolution forty years later and have a major impact on census taking.

## Revitalizing the Census: The Depression and the War Years

In the late 1920s, though, reapportionment seemed a relatively unimportant issue to most statisticians and even to many bureau officials. The advice of professional statisticians became increasingly important in the twentieth-century Census Bureau, to the point that statisticians saw themselves as a lobbying group that pressured the bureau to collect and publish this or that kind of data. Nevertheless, despite these changes, the agency and the census instrument were still political tools. Their fortunes and challenges were intimately connected to broader political and social trends in the nation. Thus, just as political historians mark the coming of the Great Depression as a major point of change in our national history, so, too, the Great Depression dramatically changed the census and the American statistical system. The 1930 census was taken less than six months after the crash of the stock market. As the depression tightened its grip on the nation, the Census Bureau found itself drawn into a contentious debate about the scope and character of unemployment and thus the causes of the economy's slide. The bureau leadership was sympathetic with President Herbert Hoover's stance toward unemployment, and thus the bureau figures from the 1930 census and a special 1931 census of unemployment put the best face on what clearly was a dramatic rise in unemployment nationally. Democratic claims that the depression was deepening were overstated, census officials claimed. Unemployment always existed—especially during the winter and early spring months when the census was taken.

As it turned out, Hoover's evaluation of the depth of the depression and the appropriate way to cope with it was soundly rejected by the voters in 1932. In the spring of 1933, Franklin Roosevelt's New Dealers came to town intent on remaking American government and society and ending the depression. The Census Bureau came in for particular scrutiny as one of Hoover's favorite agencies, and—with the exception of the Wilson years—as a longtime Republican stronghold. Director William Mott Steuart resigned and was replaced by William Lane Austin. Assistant Director Joseph Hill, who was in his seventies



and had first worked on the census in 1900, was transferred to a newly created “research” position. American Statistical Association president Stuart A. Rice replaced Hill as assistant director. Rice, a sociology professor at the University of Pennsylvania at the time, intended to stay in government for a year or so. He ultimately remained as a major figure in the federal statistical system until 1954—at the Census Bureau, then on the Central Statistical Board, and finally in the Division of Statistical Standards in the Bureau of the Budget (now the Office of Management and Budget).

Revitalizing the Census Bureau in the mid-1930s was no easy task. Throughout the 1930s, committees of statisticians, economists, and other academics—most notably the Committee on Government Statistics and Information Services—investigated bureau practices and found them wanting. Officials from Roosevelt’s alphabet agencies called upon the bureau to provide data on the socioeconomic situation of the population. Congress built the grant-in-aid system to allocate tax money from the federal to state and local governments, and they sought population data on poverty, income distribution, and migration. Yet the Census Bureau of the early 1930s had yet to devise a credible measure of unemployment, had no data on income, and did not measure poverty. The New Dealers persevered and, by the end of the decade, Congress had proposed the introduction of sample surveys to measure unemployment, reorganized the agency’s bureaucratic structure, increased the statistical training of employees, built a research unit, and analyzed and revamped classification systems. Many of these innovations became part of the 1940 census. The sixteenth census, in 1940, included the use of a sample long form for the first time, a housing census, and evaluation studies to systematically measure the level of accuracy of the enumeration, tabulation and coding procedures, coding bias, and sampling error.

Among those who worked on the probability sampling-based trial Census of Unemployment at the Census Bureau were Calvert Dedrick, Morris Hansen, Samuel Stouffer, and Frederick Stephan (Anderson 1988; Duncan and Shelton 1978). Hansen was then assigned, with a few others, to explore the field of sampling for other possible uses at the bureau and went on to work on the 1937 sample Census of Unemployment. After working on the sample component of the 1940 decennial census (under the direction of W. Edwards Deming), Hansen worked with others (for example, Jerome Cornfield, Lester Frankel, William Hurwitz, and J. Steven Stock) to redesign the

unemployment survey based on new ideas on multistage probability samples and cluster sampling (Hansen and Hurwitz 1942a and 1942b).

## The Discovery of the Undercount

It was as part of the general trend in statistical innovation at the Census Bureau that statisticians began their first systematic estimates of census undercount. Ever since George Washington had complained in the 1790s, officials had been concerned about undercounting the population. And ever since local officials had objected to their census counts in the 1870s, census officials had tried to improve the count to minimize errors. Much administrative reform of census procedures had been aimed at minimizing undercounts, overcounts, miscounts, padding, or curbstoning. But the growth of other large-scale administrative data systems—vital records and public health data, in particular—and the development of the disciplines of demography and statistics prompted a new set of questions. On the one hand, these other data systems sometimes generated alternative estimates of the population or segments of it. On the other, demographers and statisticians needed census counts for the denominators for their rate calculations or as sampling frames for other studies. Accordingly, both the professional community of statisticians and demographers and census officials began to conceive of precisely measuring the undercount for particular segments of the population.

The issue began to be framed after a somewhat serendipitous natural experiment in 1940. The selective-service registration of October 1940 allowed demographers to compare the April 1940 census counts of men of draft age (twenty-one to thirty-five) with the counts of men who registered for the draft. Daniel O. Price (1947) reported that nationally the 1940 census had underreported about 2.81 percent of the men in the age cohort (the draft registration recorded some 453,000 more men than did the census). More significant, though, was the finding that the level of the undercount varied by region and race. Some 13 percent of the Black men of draft age were missed in the census. Nationally, 229,000 more Black men registered for the draft than would have been expected from the census estimate. Price also demonstrated that the Black men registered for the draft in dramatically higher numbers in urban states than would have been expected from the April census counts. He could not, however, determine whether such men had migrated to an urban area between the time of

the census and the registration date or, rather, the census had done a poorer job of counting urban residents. Price concluded his article by discussing some of the implications of his findings, particularly that the standard indicators of vital rates, such as death rates, would have to be recalculated in light of these data.

Additional studies followed in later years. In 1955, as part of the general project to develop a definitive demographic analysis of the American population, Ansley Coale (1955), published a revision of census figures by age, sex, and color for 1950. Coale's elaborate iterative technique used age-cohort data by race and sex from the 1930, 1940, and 1950 censuses to estimate the size of each cohort for each year, adjusted by aging the population forward each decade and correcting for deaths and immigration. Coale estimated a net census undercount of about 3.6 percent. He also concluded that the census undercounted "nonwhites" dramatically—by 12 to 13 percent in 1950. Again, because his "revisions" were in service of his larger demographic project, he ended his analysis by producing revised data by race, sex, and age for use by demographers. In the 1950s, the Census Bureau also began evaluation studies of undercounts as part of its larger evaluation program. One research thrust employed the "demographic analysis" methods, as the techniques used by Coale and Price came to be called: Information from earlier censuses and other population data sources were compared with the aggregate population counts for particular cohorts of the population. A researcher might compare the data from cohorts organized by age, race, sex, and region in two sources and calculate estimates of the differences for each cohort. The corrected cohorts were then reaggregated to generate an overall estimate of the accuracy of the count.

Demographic analysis as a technique has the advantage of being relatively easy to undertake. The researcher may have to make a very large number of individual cohort analyses. A matrix based upon, for example, two sex categories, fifteen five-year age cohorts, and two racial groups (White and non-White) in two data systems generates 120 individual cohort estimates. Price's selective service/1940 census comparison generated 98 cohort estimates of undercount—that is, for the two race categories in the forty-eight states plus the District of Columbia. Nevertheless, once the appropriate data are available, the analysis is relatively straightforward. The disadvantage of demographic analysis is its inability to pinpoint exactly why the undercount (or overcount) exists. As an aggregate methodology, it cannot identify which particu-

the individuals were missed, nor can it provide more specific information on the sources of undercount beyond the information available from the original cohort variables.

To overcome these disadvantages, the Census Bureau developed new techniques, particularly the postenumeration survey (PES). After the 1950 census, the bureau undertook a sample reenumeration of the country to try to identify households missed by the enumerators, household members who were not reported within households, and other classification and categorization errors in the original enumeration. The survey used trained interviewers to improve the quality of the information reported. The bureau then matched the information from the sample survey to the original census forms and developed estimates of the quality of the original count. Analysis of the results again indicated an undercount and poorer coverage of the nonwhite population (see table 4.1). The PES results also showed an interesting pattern: the PES uncovered only 40 percent of the “net underenumeration” expected from demographic analysis. As bureau officials later explained, the PES “was very successful in finding space that the original census enumerator had missed but was much less effective in uncovering missed persons—those residing in previously enumerated space who were unreported and those without any recognizable place of residence” (Pritzker and Rothwell 1968, 61, 63).

By the time of the 1960 census, census officials and the community of professional demographers were well on their way to understanding census undercounts. They built evaluations procedures in the form of a postenumeration survey and demographic analysis into the 1960 census design. Professional discussions continued quietly on the best means to estimate undercounts for particular groups in the population, as well as to develop new census methodologies to count more accurately in the first place. These methodological discussions are evident in the Census Bureau technical reports and in the general social science literature of the late 1950s and early 1960s.

What was still missing from the undercount discussion was any sense that there were any larger political or social implications of the discovery of the undercount. From the 1940s through the mid-1960s, the literature was totally “methodological”—of interest to demographers, statisticians, and survey researchers but not to members of Congress, policy makers, or the undercounted communities. This situation was to change dramatically in the mid-1960s.

## The Politicization of the Undercount

Three separate trends merged to propel the census undercount onto the political stage in the mid-1960s. First, since the turn of the century, Congress had built the grant-in-aid system as a mechanism for allocating revenue to state and local governments. By the 1960s, Congress had used the grant-in-aid system to allocate federal funds for initiatives in vocational education, highway construction, agricultural extension, and public health programs. The extensive programs of the postwar era, for example, urban renewal, the interstate highway system, and hospital construction subsidies had used the grant-in-aid system. In 1960, 15 percent of state and local funding came from federal aid. Census data underpinned many of the formulas Congress used to allocate the funds to particular jurisdictions.

Second, in 1962, the Supreme Court ruled that malapportioned state legislatures were unconstitutional and opened the way for the decade of lawsuits that led to the “reapportionment revolution” of the 1960s. For forty years the federal courts, and particularly the Supreme Court, had refused to rule on apportionment cases. The courts had argued that legislative apportionment was strictly a legislative matter. In 1962, the Supreme Court reversed itself and ruled in *Baker v. Carr* (1962) that the Tennessee legislature had to be reapportioned. A series of cases that followed overthrew apportionments in other legislatures and in Congress. By 1964, the phrase “one man, one vote” had entered the nation’s political vocabulary to define the new principle of legislative apportionment. Suddenly, accurate census data for small geographic areas came to be of added importance. Not only were the grants-in-aid of the Great Society programs of the mid-1960s prompting increased use of census data. The apportionment cases also suggested that the Census Bureau had a strict constitutional duty, under the equal protection clause of the Fourteenth Amendment, to count everyone.

Third, the civil rights movement of the 1950s and 1960s framed many of its arguments about discrimination in the labor force, in public participation in the community, and in access to housing in terms of underrepresentation of minorities in these areas of American life. Attorneys and activists used census data to make the case. If, for example, 30 percent of a local labor market was Black, then, activists argued, true nondiscrimination would require that Blacks hold roughly 30 percent of the jobs up and down the occupational scale. After a

decade of lobbying and political action, Congress responded and passed several major pieces of civil rights legislation in the late 1950s and early 1960s, which also relied on high-quality census data for administration and implementation. For example, the Voting Rights Act of 1965 enforced the Fifteenth Amendment to the Constitution, which stated that the “right . . . to vote shall not be denied or abridged by the United States or by any State on account of race, color, or previous condition of servitude.” Congress created clear numerical tests of compliance with the constitutional goals of voting rights. As written in the original 1965 provision, if a state used a literacy test for voter registration, and if voter registration or turnout was less than 50 percent of the voting-age population of a jurisdiction, then the law presumed a violation of the Fifteenth Amendment. In such a case, the literacy tests were suspended, and the Justice Department could send federal registrars and election observers to monitor elections. Such jurisdictions also had to “preclear” any new voting qualifications with the attorney general of the United States. Six southern states came under these rules; counties in several other states were also affected. Again, census data would have a direct impact on a highly charged political issue.

During the planning phase for the 1970 census, all of these issues came into sharp focus. The differential census undercount of hard-to-count populations, especially minorities and the urban poor, took on new meaning. Once the Supreme Court had invalidated the massive legislative malapportionments of the past, the statisticians and politicians realized that the census undercount also could have the effect of denying representation to the uncoun­ted. The undercount ceased to be a technical problem of census field procedures; it became an explosive political issue. A 1967 conference on social statistics and the city marked the change (Heer 1967). At that conference, census officials and prominent social scientists met in Washington to discuss the extent and cause of the undercounts. They articulated the constitutional principle that made elimination of the differential undercount imperative. “Where a group defined by racial or ethnic terms, and concentrated in special political jurisdictions,” wrote David Heer, in the conference report, “is significantly undercounted in relation to other groups, then individual members of that group are thereby deprived of the constitutional right to equal representation in the House of Representatives and, by inference, in other legislative bodies.” They are also “deprived of their entitlement to partake in federal and other programs designed for areas and populations with their characteristics” (Heer 1967, 11). In other words, miscounting the population could



unconstitutionally deny minorities the political representation or protection guaranteed by the Voting Rights Act. It could also deny local jurisdictions grant funds from federal programs. Suddenly, Congress, demographers and statisticians, the Census Bureau, and state and local government officials recognized that they not only had to estimate the differential undercount, but they would also have to correct for it. Only slowly, however, did they realize just how hard it would be to develop the administrative and statistical innovations necessary to do so.

# Margo Anderson

## **The Census and the Japanese “Internment”: Apology and Policy in Statistical Practice**

### **PRELIMINARIES**

In March 1942, starting less than four months after the December 7, 1941, attack on Pearl Harbor, and continuing to August 1942, the US government rounded up over 100,000 West Coast Americans of Japanese ancestry and, without further investigation or due process, incarcerated them in purpose-built concentration camps, most for the duration of World War II. At the time, public officials argued—and the courts and Congress agreed—that the Japanese-ancestry population posed a security threat, and thus “military necessity” justified the mass incarceration. Small area data tabulations from the 1940 census and technical expertise from high-level census officials provided crucial administrative support for the program, so much so that the army’s *Final Report* (US Army 1943, 352) on the evacuation credited the 1940 census as the “most important single source of information prior to the evacuation” that “became the basis for the general evacuation and relocation plan.”

Scholars from the 1940s onward challenged the legitimacy of the “military necessity” justification of the program (Rostow 1945; Thomas and Nishimoto 1946; Thomas 1952, tenBroek, Barnhart, and Matson 1954; Grozdins 1949), but they didn’t find broad political sup-

port until the development of a “redress” movement within the Japanese American community in the 1970s that successfully lobbied for a formal federal re-examination (Weglyn 1976; Daniels 1993; 2013). In the early 1980s, the Commission on Wartime Relocation and Internment of Civilians investigated the program again and repudiated the rationale of military necessity, instead writing that the policy was the result of racial prejudice, war hysteria, and failure of political leadership (CWRIC 1997). In 1988, federal law called for a formal apology and authorized reparations for survivors. Presidents George H. W. Bush and Bill Clinton implemented that program. In the 1980s, the federal courts also repudiated the Supreme Court decisions, which had upheld the constitutionality of the internment during World War II (Irons 1983; 1989). Since then, an even greater outpouring of historical analysis has delved into the history of the event (see Densho: The Japanese American Legacy Project, <https://densho.org/about-densho/>).

At the time of the redress movement agitation and federal legislative activity in the 1970s and 1980s, the actions and policies of the Census Bureau and its participation in the evacuation and incarceration received relatively less attention than did the more prominent federal participants in the program, notably the Justice Department, the Army, and White House policy makers. For its part, the Census Bureau defended its actions as lawful and within the scope of the agency’s mandates.

But as with the earlier scholarly research that had provided historical documentation for the redress movement, the evidence uncovered as part of the Commission’s re-examination prompted redress activists, official statisticians, and human rights scholars to probe the Census Bureau role anew; a series of papers by me and William Seltzer delved into the Census Bureau’s activities during World War II (Seltzer and Anderson 2000; 2001). In 2000, Census Director Kenneth Prewitt responded to that historical work and issued an apology on the part of the agency (quoted in U.S. Census Bureau 2002, 16):

The historical record is clear that senior Census Bureau staff proactively cooperated with the internment, and that census tabulations were directly implicated in the denial of civil rights to citizens of the United States who happened also to be of Japanese ancestry....

I would also like to state clearly that for many years the Census Bureau was less than forthcoming in publicly acknowledging its role in the internment process. Silence was not the worst offense, for there is ample evidence that at various times the Census Bureau has described its role in such manner as to obfuscate its role in internment. Worst yet, some Census Bureau documents would lead the reader to believe that the Census Bureau behaved in a manner as to have actually protected the civil rights of Japanese Americans. This distortion of the historical record is being corrected.

The internment of Japanese Americans was a sad, shameful event in American history, for which President Clinton, on behalf of the entire federal government, has forthrightly apologized. The Census Bureau joins in that apology and acknowledges its role in the internment.

With that statement, the bureau hoped to put to rest any lingering issues about its participation in the episode, correct the historical record, and assure the public that the agency understood its role as a data steward and had procedures in place to protect census respondents from the harms of the war years. Prewitt continued, “In the post-war period, important safeguards to protect against the misuse of census tabulations have been instituted, notably stronger legal provisions to protect data confidentiality and the Census Bureau’s introduction of disclosure avoidance techniques. Adherence to these safeguards precludes a repeat of the 1941/42 behavior.” He concluded by noting the critical importance of census data for civil rights enforcement: “Over the past half-century, and especially following the

Civil Rights Act of 1964 and the Voting Rights Act of 1965, the critical importance of summary data for enforcement of voting rights and civil rights in this age is an important contrast to the misuse of information in the early 1940s.”

### **IS THIS THE END OF THE STORY?**

The short answer is no. Twenty years after his apology, Kenneth Prewitt (2019) acknowledged that the bureau’s participation in the evacuation and incarceration still plagued the agency. “The remarkable thing about what happened in 1942, basically,” he told NPR in 2019, “by using 1940 census data with the roundup of the Japanese Americans on the West Coast is that we’re still talking about it 70 years later. Every census that comes along—once again, the Japanese case gets brought to the surface. And the argument always is you can’t trust the Census Bureau. Look what they did back in 1940. So, yes, it cast a very long dark shadow.”

Census official Christa Jones made a similar evaluation in 2017: “With each decennial census for the last 40–50 years, the question of whether the Census Bureau provided confidential data to facilitate the evacuation and relocation of the Japanese Americans during World War II (i.e., the “Japanese internment”) rematerializes, seemingly never resolved.”

Yet, as Ruggles and Magnuson (2020) have recently argued, this concern about misuse of the census data lingers despite the fact that “there are no documented instances in which the identity of anyone in the decennial census has been determined by anyone outside the Census Bureau.” What the Census Bureau calls disclosure control protects individual level responses from being reported in any data publication. They emphasize that “disclosure control for population census publications was introduced in 1929 and is unchanged since then.”

So the question is why, despite the successful disclosure control methods for microdata, and the apology from the Census Bureau for its participation in the evacuation and incarceration of Japanese Americans during World War II, the issue is still not “resolved” (using Jones’s term). Why weren’t the apology and policy changes enough?

One answer, hinted at in Prewitt's comments, is the debate in the larger political environment about the use and abuse of government data in general, and the digital revolution that has both democratized data use and heightened the public's awareness of the importance of privacy, data collection and management, and the uses and interpretation of data.

But a second and less understood answer may be in the way that public concerns about "privacy in the digital age" and what statisticians call "statistical confidentiality" developed historically and intersected in law and administrative practice, influencing both the Census Bureau's behavior during World War II and its understanding of its responsibilities afterwards. And for that perspective, it's useful to revisit the 1940s experience in more detail.

## **CENSUS LAW**

The 1940 census was governed by the 1929 omnibus census statute, the Reapportionment Act of 1929 (P.L. 71-12). Two provisions of that statute governed the release of data on individuals. The first, section 11, affirmed that census data were to be used for "statistical purposes" only and that "No publication shall be made by the Census Office whereby the data furnished by any particular establishment or individual can be identified, nor shall the Director of the Census permit anyone other than the sworn employees of the Census Office to examine the individual reports." The second provision, section 18, somewhat contradictorily provided that the director was "authorized, at his discretion," "to furnish" governors, courts of record, or individuals "such data from the population schedules as may be desired for genealogical or other proper purposes," with the caveat that "in no case shall information furnished under the authority of this Act be used to the detriment of the person or persons to [whom] such information relates."

It was the ambiguity in the relationship between those two provisions that critics of the bureau's actions raised when challenging the census role in the evacuations. The issue was framed during the war, as well as in the 1980s, by the redress movement and in CWRIC



hearings and deliberations. During the war, in late March 1942, Congress temporarily overrode the requirements of section 11, in Section 1402 of the Second War Powers Act (P.L. 77-507), which provided that

notwithstanding any other provision of law, any record, schedule, report, or return, or any information or data contained therein, now or hereafter in the possession of the Department of Commerce, or any bureau or division thereof, may be made available by the Secretary of Commerce to any branch or agency of the Government, the head of which shall have made written request therefor for use in connection with the conduct of the war.

That authorization came too late for use during the evacuation of the Japanese-ancestry population already in progress, but it was used for other microdata releases from 1942 to 1947. The bureau has acknowledged the legislative change during the war, emphasizing its 1947 repeal, and the restoration of the previous standard (Jones 2017).

The historical record on the meaning of section 18 of the 1929 census act is more confusing in terms of when and how that provision was added to the census statute, whether it was deployed to determine any particular data releases, or whether respondents ever sued the Census Bureau, citing section 18. This confusion became particularly evident when the redress movement challenged the bureau to acknowledge that it had violated the “detriment” clause of section 18 through its participation in the evacuation administration.

The issue emerged publicly as Census Director Vincent Barabba promoted census participation in the upcoming census in early 1980. The trigger was a January 25, 1980, front-page article in the *Pacific Citizen*, the “national publication of the Japanese Americans Citizen League” [JACL], on the Census Bureau director’s press conference explaining the upcoming census. “Census Bureau refused to let War Dept. have Nikkei names” was the article’s headline. The piece continued:

Those who cooperate with the population count on April, this year, will be guaranteed their privacy would be protected, even the illegal immigrant, assured Vincent P. Barabba, director of the U.S. Census Bureau. ... He pledged the census would not be turned over to any law enforcement agency. As historical precedent, Barabba cited the Census Bureau's refusal in 1941 to give the names of Japanese Americans to the War Department despite the anti-Japanese hysteria that struck after Pearl Harbor.

By the spring of 1980, activists involved in the redress campaign challenged that claim and initiated efforts to press the bureau to rediscover and explain its role in the program.

Raymond Okamura, a former Gila River evacuee and longtime JACL activist, made the first approach to the bureau in late May 1980, after he had received a copy of General DeWitt's *Final Report* from Michi Weglyn. That report provided a detailed accounting of the Census Bureau's role in the evacuation program and credited the data from the 1940 census as the "most important single source of information prior to the evacuation." Referencing Barabba's confidentiality claim, Okamura pressed the bureau: "The Bureau of Census has been less than honest about this matter. I believe some sort of apology or explanation is in order for disseminating false and misleading information." In a letter to Dr. Mary Watanabe, "JACL Census Liaison and Census Advisory Committee Member," Okamura was blunter: "I was foolish in believing the Census Bureau's propaganda regarding confidentiality of records. Government agencies simply cannot be trusted, and I am embarrassed to have been so naïve."<sup>1</sup>

Bureau officials did not respond promptly to Okamura, perhaps because Okamura mistakenly sent his original inquiry to former Census Director Manuel Plotkin rather than then current director, Vincent Barabba. But the Asian stakeholder community had been alerted, and on June 9, 1980, Christopher Chow, a census information officer in San Francisco, wrote to Barabba saying he was getting calls

on the issue and urged the bureau to respond: “I think the credibility of the Bureau with regard to confidentiality could be seriously jeopardized if this issue is not clarified soon—before someone makes it (and our reference to it) a public issue.” He recommended “that the bureau review this matter thoroughly; that the Bureau undertake to explain what kind of information was on those punch cards, precisely how the data was used ..., and what procedures are now in place as safeguards, and what policies would apply in such a situation were it to recur in the future .... This issue should be cleared up once and for all.”<sup>2</sup>

The bureau acknowledged Okamura’s inquiry on July 2, indicating that “since your inquiry refers to historical information, additional time is required to complete our research.” Theodore (Ted) Clemence, senior advisor in the bureau, researched the issues. The first draft response in July, “Confidentiality of Census Records during World War II,” drew a scathing response from Christopher Chow. “The Census Bureau will be blown right out of the water if you release the draft public statement.” Chow sent a four-page critique, challenging the defensiveness of the draft, noting its omissions and its silences on the clear harms that the evacuation and incarceration inflicted on Japanese Americans. Clemence went back to the drawing board, acknowledging in an October 1980 memo to Barabba that he had learned since August that the Second War Powers Act had suspended confidentiality “temporarily.” He told Barabba he had received a phone call from Mary Watanabe of the Advisory Committee and that “we should be prepared for this topic to be brought up during the committee mtg. on October 30.” He also had tried to “reach” Calvert Dedrick, the former census official who was deployed to the Army in February 1942 and who oversaw the Census Bureau’s participation in the program. Dedrick had retired in 1968; Clemence told Barabba he couldn’t reach him.<sup>3</sup>

On December 16, 1980, Barabba provided a six-page, single-spaced letter to Okamura responding to his questions. He claimed that “as to the 1980 census publicity about the Bureau’s refusal to

disclose names to the War Department in 1942, the statements were made in good faith based upon existing knowledge of the events.” He continued, “Based on our recent review of the events, it is apparent that a different account would be appropriate in the future,” but still claimed that the “dissemination of the 1940 census statistics to assist the war agencies in 1942 did not, in our opinion, violate the applicable census law, *even though the aggregated information may have been used to the detriment of groups or individuals*” [emphasis added].<sup>4</sup>

On January 23, 1981, Raymond Okamura published an article in *Pacific Citizen* on Barabba’s response. With the headline “Census data used to plan ‘42 Evacuation,” Okamura’s article quoted from Barabba’s response, and added detailed discussion of the Second War Powers Act and the intent of Congress to repeal census confidentiality as a war measure. Okamura warned that the repeal policy had never been challenged in court, and alerted readers to the lingering dangers that another national emergency could lead Congress to do it again. The *Pacific Citizen* editor also reported that the recent JACL national convention had recommended that the organization “study” these matters, including whether JACL should “initiate legislation to strengthen census confidentiality” and “urge the Redress Commission to investigate the role of the Census Bureau in 1942.” Okamura sent Barabba’s letter to the *Washington Post* as well, and the *Post* reported on it on February 15, 1981. By that time, Barabba had resigned with the change of administration and the bureau was without a permanent director until his successor, Bruce Chapman, took up his post in the summer of 1981.<sup>5</sup>

In February 1981, Okamura contacted the bureau again, asking Acting Census Director Daniel Levine to send an enclosed letter to Calvert Dedrick, who had been mentioned in Barabba’s December letter to Okamura. Okamura asked Dedrick for further information about his role in the evacuation. Dedrick replied to Okamura on March 12, providing details on his role in a three-page, single-spaced letter, though he also claimed that his memory of the events had faded. He concluded with a spirited defense of bureau actions: “I hope that this

letter will be helpful in your research effort. I also hope that you will help to kill the canard which has been spread about the violation of its confidentiality laws and rules by the Bureau of the Census.”<sup>6</sup>

Okamura was not satisfied. In September 1981, in testimony before the Commission on Wartime Relocation and Internment of Civilians, Okamura argued that census confidentiality was a “myth” (Okamura 1981). He recounted his prior research on the bureau’s claims, and the inadequate historical understanding of the impact of the Second War Powers Act, and called upon the Commission to “further investigate the role of the Census Bureau in the mass incarceration of Japanese Americans, and to make appropriate recommendations for legislation to strengthen the census confidentiality law.”

Bureau officials thus consistently denied that their actions constituted a violation of section 18. Perhaps the most pointed example of the differing positions came in the November 1981 questioning by the Commission on Wartime Relocation and Internment of Civilians committee member Judge William J. Marutani of Calvert Dedrick (1981), the census assistant director for statistical research in 1942. Dedrick oversaw the statistical work and administration of the evacuation for the Western Defense Command in 1942 and 1943.

JUDGE MARUTANI: So what I understand you are saying is you did know this information was going to be used by the WCCA [Wartime Civil Control Administration] to round up persons of Japanese ancestry so that they could be sent to the camps?

MR. DEDRICK: Well, that was my duty to try to use it.

JUDGE MARUTANI: And that was the reason for why this information was released, and you were aware of it?

MR. DEDRICK: Yes.

JUDGE MARUTANI: What about the provision in the law, sir, that was existing at that time reading as follows: “That ... [and it does not say names or addresses] ... in no case shall information furnished under authority of this chapter be used to the detriment of the person or persons to [whom] such information relates.”

Dedrick responded that bureau attorneys understood the provision Marutani quoted to refer to requests from an individual for an age search, or a request for a special tabulation on an individual. He explained that an individual could request their own information “in lieu of a birth certificate” but that third parties could not receive it. Marutani followed up:

JUDGE MARUTANI: Well, as I understand it then your solicitors have advised you what this language means.... If that means that A wishes to sue B, you can’t get that information, but on the other hand, if X wishes to incarcerate 112,000 people, you can get the information, is that what your solicitors gave you?

MR. DEDRICK: ... when we make a special tabulation we try to protect confidentiality showing inclination [sic; this is a transcript error and should be “information”] about a given individual. However, the fact that there are so many Cubans living in Miami, it is a census fact, and a statistical fact, and if something happens to Cubans in Miami, that is not our fault, that is not used to their detriment. Otherwise the hands of the government of the United States, which takes its census and is responsible for the census, will be completely tied in ever ever [sic] using information from the census, because somebody might consider it to their detriment.



JUDGE MARUTANI: Well, I don't want to prolong it, I think I have the answer, but I cannot help but make one comment, that perhaps if I was a Cuban American and the census information came around, and particularly if I lived in South Florida, I may leave that space blank or put something else other than Cuban with this type of interpretation, I must tell you that.

MR. DEDRICK: Then, sir, you would be in violation of federal law.

JUDGE MARUTANI: Well, it's better than being jailed for certain.

In other words, the bureau claimed its legal responsibility was to keep individual-level data confidential. It did not extend to data tabulations used to the “detriment” of a group of respondents. Thus, there was no reason to examine the role of the bureau in the evacuation further, or apologize for its provision of the small area tabulations that were used to administer the evacuation, or to revise the census statute to prevent such tabulations from being used to the detriment of a group.

That interpretation of the census statute and the resulting single-minded focus on the need to protect individual-level data from intentional or inadvertent release continue to define the Census Bureau's approach to disclosure. It is also the key to understanding why the Prewitt apology did not “resolve” the issue.

## **DATA TABULATIONS AND THE WAR ON TERROR**

Our research (Seltzer and Anderson 2000) on the role of the census in the Japanese incarceration and evacuation was first presented in spring 2000 as part of a larger academic literature on ethics in the statistical profession and the role of official statistics in the abuse of or support for human rights. Prewitt's apology also dates from spring

2000. In 2001, the issues took on added urgency as the 9/11 terrorist attacks in the United States unleashed the “war on terror” and fear of further attacks quite reminiscent of the fear that gripped the United States in the early months after the Pearl Harbor attack. Would Muslims, Arabs, or Americans of Middle Eastern ancestry face the same kind of private and state-sponsored violence and discrimination that the Japanese-ancestry population faced in World War II?

History does matter here, and the Japanese American community played a strong role in avoiding that historical recurrence, though there were scattered incidents of private citizens attacking people they blamed for the attacks. The record is still likely more complicated than we know, but it’s clear the Census Bureau resisted providing microdata to surveillance agencies or offering to participate actively in the “war on terror.” The only mention, for example, of changes in statistical confidentiality law in the Patriot Act affected the National Center for Education Statistics, where Congress authorized the release of microdata aimed at identifying students in flight schools to surveillance agencies (Seltzer and Anderson 2002).

But what of tabulated data, that is, the kind of small area data releases that Dedrick claimed were not prohibited by the detriment clause, even if they would lead to “harm” of individuals or groups?

Unfortunately, just such a situation emerged in 2004, as Freedom of Information Act (FOIA) filings revealed that in 2002 and 2003, the Census had quietly provided the Department of Homeland Security with small area tabulations on “Arab Americans” by zip code from the 2000 census long-form sample data. As the incident became public in national media in the summer of 2004, the same conflict about “detriment” and outrage from civil rights leaders emerged in the pages of the *New York Times* (Clemetson 2004) and, somewhat later, in academic journals (e.g., El Badry and Swanson 2007).

James Zogby, the president of the Arab American Institute, raised the “trust” question and reacted in language reminiscent of the concerns of Japanese American redress activists. “As this gets out,” he noted, “any effort to encourage people to full compliance with the

census is down the tubes” (Clemetson 2004). He asked rhetorically, “how can you get people to comply when they believe that by complying they put at risk their personal and family security?”

Once again, the bureau initially defended its policies and saw no need to apologize for the data releases. Census Bureau Deputy Director Hermann Habermann responded for the bureau: “We are required to provide information to other federal agencies.... This is not a cabal calculating secret tabulations.”

The *Times* reporting continued:

Ultimately, Mr. Habermann said, any discussion about the controversial nature of the information sharing is separate from the agency’s mandate to provide information.

“The only way we can guarantee that no one will ever be harmed by our information is to release nothing,” he said. “We understand that groups can be affected by what we give out, and we understand that can be sensitive. But that is a societal debate, not a census debate.”

To its credit, the bureau did implement changes in its procedures for providing special tabulations, including review of any requests on “sensitive” populations, instructions to a requester on how they could access the data themselves, and notice that the agency would keep a publicly available list of all “custom” tabulations and that such tabulations would be available to other users upon request.<sup>7</sup>

On starting a conversation on the larger ethical and legal issues of small area census tabulations, however, bureau leadership once again argued that they were aware of the issues and had made appropriate policy changes. Deputy Director Habermann took the lead and articulated the bureau position in a series of presentations and papers. They culminated in 55 pages of papers, responses, and rejoinder in the *Journal of Official Statistics* (JOS) in 2006 (Habermann 2006a; 2006b; Fienberg 2006; Prewitt 2006; Holt 2006; Clark 2006; Anderson and Seltzer 2006; Trewin 2006).

Habermann (2006a) echoed the defenses of bureau policies that Dedrick used in the early 1980s: “The confidentiality provisions of the U.S. Census Bureau’s law, however, are not intended to prohibit the use of aggregate statistical information—uses that may result in harm to particular groups within society.” He further explained, asserting that “the U.S. Census Bureau had never extended its promise or the protections beyond individual records and information,” and that “clearly the U.S. Census Bureau has no mandate or authority to decide which users should receive or not receive aggregate data. In fact, if such a mandate were proposed, we would argue against it. It would not be suitable for the U.S. Census Bureau to exercise such authority and become a gatekeeper, determining who and what organizations could access statistical data.” He asked rhetorically: “What criteria would the U.S. Census Bureau use in its role as gatekeeper? Should the U.S. Census Bureau decide, for example, that an organization sponsored by one religion is to get information while a different one is not? How would the Bureau decide if one group was of sufficient sensitivity?” He acknowledged the critics of the Arab American tabulations release:

We understood those who wanted and/or expected the U.S. Census Bureau to prevent any possible misuse of data and to ensure that the “bad guys” did not get information that could be used to harm groups. Yet, the U.S. Census Bureau does not have a mandate to accomplish this, nor the ability to anticipate misuse. Groups are affected by the release of statistical data, and it is not possible for the U.S. Census Bureau to prevent such effects. In fact, the same statistics are often cited for and against the “interests” of certain groups. For instance, the data requested by Customs and Border Protection could also be used for civil rights enforcement or by companies seeking potential marketing opportunities.

He summarized:

What, then, is reasonable to expect from the U.S. Census Bureau? Society can expect that the U.S. Census Bureau will not release his or her individually identifiable data, and the U.S. Census Bureau will only use individual information to prepare aggregate data products. A respondent can also expect that these products will cause actions at the national and subnational level. The U.S. Census Bureau is a statistical agency and its mission is to release data, provided those data do not violate the fundamental mandate to protect an individual's information. The U.S. Census Bureau devotes considerable resources through its disclosure avoidance program to ensure that individual information is not released. Of course, the only way to guarantee—in the absolute sense—that no individual information is ever released is to provide no statistical aggregates.

The critics of Habermann's position were sharp, noting that there were both policy changes and political considerations that he needed to recognize. In his rejoinder to these critiques (Habermann 2006b), he acknowledged the "release of information about Arab Americans has caused us to reexamine the nature of our dialogue with respect to our dissemination authorities and responsibilities." He explained the policy changes that the bureau had already implemented, and acknowledged that "beyond the laws and regulations, however, leaders of statistical agencies are still moral beings and cannot ultimately use the law as an inviolable shield. They may have to make decisions about dissemination of data that could cause mission or group harm on ethical and not merely legal considerations."

Former director Kenneth Prewitt (2006) had taken Habermann to task on his treatment of what Prewitt called "mission harm," namely that incidents like the Arab American tabulations release and its aftermath not only raised the question of harm to respondents

but also damaged the Census Bureau itself and its capacity to fulfill its mission to collect and disseminate data accurately and effectively. Habermann (2006b) acknowledged the issue. He noted that “Trewin and Prewitt raise the issue of mission harm and I agree that this issue needs further discussion.” He referenced “Ken Prewitt’s description of the context in which his apology was issued.” Prewitt had noted that “the privacy debate would depress cooperation and thus an accurate count of America’s racial minorities. Some of these groups were focused on how the Japanese-American relocation in 1942 was echoing across the media coverage, knowing that their constituencies were already wary about government misuse of census answers.” He conceded: “While it is difficult to quantify the risk of mission harm, I can conceive of circumstances when the Director of the U.S. Census Bureau would refuse to release nonidentifiable information because he or she believed the public controversy resulting from that release would endanger the U.S. Census Bureau’s ability to conduct a census.”

Habermann (2006b) also responded to his critics’ proposals that some sensitive data should not be collected in the first place. He asserted that “before collecting such [sensitive] information we must engage in an open dialogue about the possible misuse of data by society.... We may want to sharpen our dialogue so that questions of possible misuse are considered by society, e.g., Congress and special interest groups, before the actual data collection.” Tim Holt (2006) noted that the technological advances in database availability and analysis raised a related threat of “third party” confidentiality breaches, and suggested that responsibility for such a breach should fall upon the user, not the Census Bureau. Habermann (2006b) expressed doubts about the possibility of holding a user “responsible” for misuse, noting that “the size of these databases and the capability to use them to breach confidentiality is largely unknown.” But he added, “it is always easier to criticize an idea than to create one, and his idea deserves further study.”

## RECENT ISSUES

The Arab American tabulations controversy both resurrected the questions about the bureau's actions during World War II and raised a new series of issues about the integrity and responsibilities of the statistical system in providing usable data for public use.

In the years that followed the *JOS* papers, bureau officials continued to focus on the issues they could most readily control, the policies in the agency to protect from “mission harm,” ethical training for staff, and particularly technical exploration of the potential for the digital revolution to upend traditional notions of privacy protection and confidentiality breaches.

The integration of perspectives of the user community in shaping new approaches to small area statistics and nondisclosure has been harder. As Prewitt noted in his 2000 apology, small area data is essential for civil rights and voting rights enforcement, and users by and large have gotten increasingly detailed tabulations and microdata samples. Those releases were protected from microdata breaches with disclosure control techniques, such as top coding, cell suppression, and swapping. But as techniques of database reconstruction, reidentification, and matching capabilities have improved, so has the potential for users to use census tabulations to identify individual respondents. As Habermann (2006b) noted, “there is pressure, at least in the United States, for smaller and smaller population groups to demand that the Census Bureau collect information about their demographic and economic characteristics. Before collecting such information we must engage in an open dialogue about the possible misuse of data by society; the best option may be to not collect the data at all.”

Not surprisingly, the increasingly large and powerful census user community, from the redistricters to the academic researchers to the state and local officials who used the data for their policy guidance, all wanted the data they were already getting—and clamored for more. In response, the bureau worked on initiatives to introduce the technique called “differential privacy” to protect the small area tabulations so they could meet the microdata protection standard of



the census statute, Title 13 of the U.S. code. In March 2018, the Trump administration demanded a question on citizenship be included on the 2020 census form and called for publication of “block-level data on the citizen voting-age population by race and ethnicity” from the 2020 census. Bureau officials responded in the fall of 2018 with the announcement of the implementation of “differential privacy” as the new data protection standard for 2020 census tabulations. “The key insight from computer science,” wrote Chief Scientist John Abowd and collaborator Ian Schmutte (2018), “is that publication of summary statistics leaks the same kind of private information as a breach. Differential privacy has emerged as a focal paradigm because it can provably circumscribe such leakages.”

And once again, the World War II experience became part of the conversation, this time in Abowd and Schmutte’s defense of the need to change the disclosure avoidance protocols for tabulations:

For an example of the harms of breaching data security, one need look no further than the Census Bureau’s activities in WWII, releasing small-area data for the purposes of Japanese internment to the Army and providing individual records of the Japanese Americans in Washington, DC to the Secret Service for purposes of surveillance .... Statutory barriers now prevent explicit data-sharing of this sort, and the U.S. Census Bureau staunchly guards those barriers.... The Secretary of Commerce’s March 26, 2018 direction to include a question on citizenship on the 2020 Census in support of providing block-level data on the citizen voting-age population by race and ethnicity makes the question of how best to protect the confidentiality of the micro-data in these publications even more salient... There remains a threat that detailed data on a sensitive population could be accidentally shared by publishing so much summary information that the underlying data can be reverse-engineered. Therefore, statistical publications should guard against database reconstruction.

## THE UNFINISHED CONVERSATION

So, what lessons might one draw from this recounting of the experience of the Census Bureau apology for this “shameful event in American history,” as Prewitt (US Census Bureau 2002) called it? The 2000 apology came more than a decade after the CWRIC report and official US government apology and compensation. It started an internal conversation within the Census Bureau and the federal statistical system about what happened during World War II, and more broadly about the harms that the work in collecting and publishing information on the population could have. It reinforced the responsibility of the agencies to enforce the confidentiality of individual responses to government surveys in law and administrative practice.

The apology did not, however, completely heal the wounds of mistrust that the public, particularly in vulnerable minority communities, expressed about future misuse of government data for discrimination, racial targeting, or law enforcement. Those fears of misuse remained latent, ready to re-emerge from new incidents of possible data harm, as with the Arab American tabulation controversy after 9/11, or the effort by the Trump administration to put a citizenship question on the 2020 census. Census officials maintained that federal law did not extend protection of individuals or groups harmed by publicly available statistical tabulations. The bureau acknowledged that such mistrust could result in the public refusing to cooperate with government surveys, and was willing to consider how such mistrust could create “mission harm” that could damage the federal statistical system itself.

To date, however, that acknowledgment has not led to renewed discussion of the implications of the “detriment” clause of Title 13. Instead, the bureau is relying on its disclosure avoidance methods and the technical innovation of differential privacy to protect its tabulations from microdata reidentification. As of this writing, the new disclosure avoidance methodology is facing major challenges. The user community has identified serious problems of accuracy in the differentially private tabulations (Macagnone 2020), which raise a

new form of “harm” in the form of flawed data. A detailed analysis of this latest controversy is beyond the scope of this paper, but it is clear that once again, the Census Bureau and the statistical user community are being called upon to reckon with past policies, the meaning of Title 13, and the role of census data in the American state.

*POSTSCRIPT, FEBRUARY 2021: THE SUPREME COURT STOPPED DONALD TRUMP’S efforts to put a citizenship question on the 2020 census form in June 2019. Trump’s efforts to mandate that the Census Bureau find other data sources to produce population counts by citizenship status also failed. Trump left the presidency on January 20, 2021. On President Joe Biden’s first day in office, he revoked the July 2019 Trump administration Executive Order mandating the Census Bureau to collect and publish tabulations on citizenship status in the 2020 census redistricting files, and the July 2020 President Memorandum mandating that the Census Bureau produce counts of “illegal aliens” to be removed from the 2020 state level apportionment counts (White House 2021). Because of the disruption of the pandemic, 2020 census results are not expected until April 2021.*

## NOTES

1. The correspondence among Okamura, Watanabe, and bureau officials reached the bureau and were archived in Recent Acquisitions, Entry 362L, Office of the Director Dr. John G. Keane (1984–89), Box 18, Folder, “Japanese Americans (WWII),” Record Group 29, National Archives and Records Administration, Washington, D.C.
2. To Vincent Barabba, June 9, 1980. See note 1 above.
3. Clemence, see note 1 above. For Clemence’s career, see his 1988 obituary in the *Washington Post*, [www.washingtonpost.com/archive/local/1988/10/12/obituaries/263a965d-8492-4efe-9079-668f04388ec7/](http://www.washingtonpost.com/archive/local/1988/10/12/obituaries/263a965d-8492-4efe-9079-668f04388ec7/).
4. Barabba, June 9, 1980. See note 1 above.
5. *Pacific Citizen*, January 23, 1981. See, for example, Spencer Rich, “Census Bureau Aided ‘42 Roundup of Nisei,” *Washington Post*, Feb 15, 1981, G2.

6. Okamura to Dedrick, February 25, 1981; Dedrick to Okamura, March 12, 1981; see note 1 above.
7. Detail on the evolution of the procedures in 2004 and 2005 is available in U.S. Census Bureau, *History: 2000 Census of Population and Housing*, vol. 2 (Washington, D.C.: GPO, 2009), 433–34, [www.census.gov/history/pdf/Census2000v2.pdf](http://www.census.gov/history/pdf/Census2000v2.pdf).

## REFERENCES

- Abowd, John M., and Ian M. Schmutte. 2018. "An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices." August 21, 2018. Forthcoming in *American Economic Review*. <https://arxiv.org/pdf/1808.06303.pdf>.
- Anderson, Margo, and William Seltzer. 2006. "Discussion." *Journal of Official Statistics* 22 (4): 641–49.
- Clark, Cynthia Z. F. 2006. "Discussion." *Journal of Official Statistics* 22 (4): 637–39.
- Clemetson, Lynette. 2004. "Homeland Security Given Data on Arab-Americans." *New York Times*. July 30. [www.nytimes.com/2004/07/30/us/homeland-security-given-data-on-arab-americans.html](http://www.nytimes.com/2004/07/30/us/homeland-security-given-data-on-arab-americans.html).
- Daniels, Roger. 1993. *Prisoners Without Trial: Japanese Americans in World War II*. New York: Hill and Wang.
- . 2013. *The Japanese American Cases: The Rule of Law in Time of War*. Lawrence: U. Kansas Press.
- Dedrick, Calvert L. 1981. "Testimony before CWRIC, November 3, 1981." Record Group 220, Records of Temporary Committees, Commissions, and Boards, 1893–2008, Commission on Wartime Relocation and Internment of Civilians. 7/31/1980-6/30/1983. National Archives Records Administration, College Park, MD, pp. 170–90.
- El-Badry, Samia, and David Swanson. 2007. "Providing Census Tabulations to Government Security Agencies in the United States: The Case of Arab Americans." *Government Information Quarterly*. 24: 470–87.
- Fienberg, Stephen E. 2006. "Discussion." *Journal of Official Statistics*. 22 (4): 615–19.

- Grodzins, Morton. 1949. *Americans Betrayed: Politics and the Japanese Evacuation*. Chicago: U. Chicago Press.
- Habermann, Hermann. 2006a. "Ethics, Confidentiality, and Data Dissemination." *Journal of Official Statistics*. 22 (4): 599–614.
- . 2006b. "Rejoinder." *Journal of Official Statistics*. 22 (4): 651–53.
- Holt, Tim. 2006. "Discussion." *Journal of Official Statistics*. 22 (4): 627–29.
- Irons, Peter. 1983. *Justice at War: The Story of the Japanese American Internment Cases*. New York: Oxford University Press.
- . 1989. *Justice Delayed: The Record of the Japanese American Internment Cases*. Middletown, CT: Wesleyan U. Press.
- Jones, Christa. 2017. "Documents and Citations": Original Sources and Research Concerning Census Bureau Efforts to Support Japanese Internment, March 22, 2017, available at <https://ecommons.cornell.edu/bitstream/handle/1813/66186/2017-005%20May-22%20MEMO%20Jones%20Files.pdf?sequence=3>.
- Macagnone, Michael. 2020. "Researchers Push for Public Test of Census Privacy Tools." *Roll Call*, June 10. <https://www.rollcall.com/2020/06/10/researchers-push-for-public-test-of-census-privacy-tools/>.
- Okamura, Raymond Y. 1981. "The Myth of Census Confidentiality." *Amerasia Journal*. 8 (2): 111–20.
- Prewitt, Kenneth. 2006. "Discussion: Statistics in the National Interest." *Journal of Official Statistics*. Vol. 22, No. 4, 621–25.
- . 2019. Interview on All Things Considered, July 6. [www.npr.org/2019/07/06/739227590/former-census-director-citizenship-question-to-hurt-2020-accuracy](http://www.npr.org/2019/07/06/739227590/former-census-director-citizenship-question-to-hurt-2020-accuracy).
- Rostow, Eugene. 1945. "The Japanese American Cases: A Disaster." *Yale Law Journal*. 54(3): 459–533.
- Ruggles, Steven, and Diana L. Magnuson. 2020. "'It's None of Their Damn Business': Privacy and Disclosure Control in the U.S. Census, 1790–2020." Presentation at the annual meeting of the Population Association of America, April 2020.
- Seltzer, William, and Margo Anderson. 2000. "After Pearl Harbor: The Proper Role of Population Data Systems in Time of War." Paper

- presented at the Annual Meeting of the Population Association of America, Los Angeles, March.
- . 2001. “The Dark Side of Numbers: The Role of Population Data Systems in Human Rights Abuses,” *Social Research* 68 (2) Summer: 481–513.
- . 2002. “NCES and the Patriot Act,” Paper presented at the Joint Statistical Meetings, New York, August.
- tenBroek, Jacobus, Edward N. Barnhart, and Floyd Matson. 1954. *Prejudice, War, and the Constitution*. Berkeley: U. California Press.
- Thomas, Dorothy S. *The Salvage*. 1952. Berkeley: U. California Press.
- Thomas, Dorothy S., and Richard Nishimoto. [1946] 1969. *The Spoilage*. Berkeley: U. California Press.
- Trewin, Dennis. 2006. “Discussion.” *Journal of Official Statistics*. 22 (4) 631–35.
- US Army. Western Defense Command and Fourth Army. 1943. *Final Report: Japanese Evacuation from the West Coast, 1942*. Washington, DC: US Government Printing Office.
- U.S. Census Bureau. 2002. “Census Confidentiality and Privacy: 1790–2002.” [www2.census.gov/library/publications/2003/comm/monograph-confidentiality-privacy.pdf](http://www2.census.gov/library/publications/2003/comm/monograph-confidentiality-privacy.pdf).
- U.S. Commission on Wartime Relocation and Internment of Civilians (CWRIC). 1997. *Personal Justice Denied: Report of the Commission on Wartime Relocation and Internment of Civilians*. Seattle: U. Washington Press. Reissued with a new forward by Tetsuden Kashima, 1982. Washington DC: US Government Printing Office. CWRIC.
- Weglyn, Michi Nishiura. 1976. *Years of Infamy: The Untold Story of America’s Concentration Camps*. New York: William Morrow and Co.
- White House. 2021. Executive Order on Ensuring a Lawful and Accurate Enumeration and Apportionment Pursuant to the Decennial Census. Presidential Actions. January 20. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-ensuring-a-lawful-and-accurate-enumeration-and-apportionment-pursuant-to-decennial-census/>.

# Reading Unit 3:

# Data Visualization

(Estimated Reading Time: 1.5 hours,  
Complete Reading by 01/31/2022)

**UO31:** Through readings, group discussion and reflection, students will analyze how a statistician's social identity translates into biased understanding of the history of statistical thought. (CO 11)

**UO32:** Through readings, group discussion and reflection, students will critique the dominant narrative about the history of data visualization by comparing and contrasting euro-centric analysis of the history of statistical thought and more diverse understandings of the history of statistical thought. (CO 11)

1. Friendly, M. (2008). A brief history of data visualization. In *Handbook of data visualization* (pp. 15-56). Springer, Berlin, Heidelberg.
2. Lundeen, P. (July 22 2021). REVIEW: Data + Diversity-Exploring the Data Visualizations of W.E.B. Du Bois. <https://nightingaledvs.com/review-data-diversity-exploring-the-data-visualizations-of-w-e-b-du-bois/> (accessed 2022-01-08)).
3. Evergreen, S. (July 19 2019). Beyond Nightingale: Being a Woman in Data Visualization. <https://nightingaledvs.com/beyond-nightingale-being-a-woman-in-data-visualization/> (accessed 2022-01-08).
4. Ying, L. P. (August 25 2021). Asia's data scene deserves greater attention. That's why we are starting a movement. <https://nightingaledvs.com/asias-data-scene-deserves-greater-attention/> (accessed 2022-01-08).



# Reading Unit 3:

## Discussion Questions

1. In the first reading, the history of data visualization is given. How much of the history of data visualization is centered in Western Europe and the United States?
2. The last three readings explore experiences of scientists that are not white and male. How does each of these scientists experience tokenism or structural violence in their career? How are they resisting the cultural norm?

# Reading Unit 3:

## Reflection Questions

1. What are some of the ways that a Euro-centric focus on the history of data visualization affects and distorts our conceptualization of data visualization today?
2. What are some of the ways that modern statisticians are working to rewrite the history of data visualization and make it more inclusive?

# A brief history of data visualization

Michael Friendly\*

March 21, 2006

## Abstract

It is common to think of statistical graphics and data visualization as relatively modern developments in statistics. In fact, the graphic representation of quantitative information has deep roots. These roots reach into the histories of the earliest map-making and visual depiction, and later into thematic cartography, statistics and statistical graphics, medicine, and other fields. Along the way, developments in technologies (printing, reproduction) mathematical theory and practice, and empirical observation and recording, enabled the wider use of graphics and new advances in form and content.

This chapter provides an overview of the intellectual history of data visualization from medieval to modern times, describing and illustrating some significant advances along the way. It is based on a project, called the *Milestones Project*, to collect, catalog and document in one place the important developments in a wide range of areas and fields that led to modern data visualization. This effort has suggested some questions of the use of present-day methods to analyze and understand this history, that I discuss under the rubric of “statistical historiography.”

## 1 Introduction

*The only new thing in the world is the history you don't know.* —Harry S Truman

It is common to think of statistical graphics and data visualization as relatively modern developments in statistics. In fact, the graphic portrayal of quantitative information has deep roots. These roots reach into the histories of the earliest map-making and visual depiction, and later into thematic cartography, statistics and statistical graphics, with applications and innovations in many fields of medicine and science that are often intertwined with each other. They also connect with the rise of statistical thinking and widespread data collection for planning and commerce up through the 19<sup>th</sup> century. Along the way, a variety of advancements contributed to the widespread use of data visualization today. These include technologies for drawing and reproducing images, advances in mathematics and statistics, and new developments in data collection, empirical observation and recording.

From above ground, we can see the current fruit and anticipate future growth; we must look below to understand their germination. Yet the great variety of roots and nutrients across these

---

\*This paper was prepared for the *Handbook of Computational Statistics on Data Visualization*. Michael Friendly is Professor, Psychology Department, York University, Toronto, ON, M3J 1P3 Canada, E-mail: friendly@yorku.ca. This work is supported by Grant 8150 from the National Sciences and Engineering Research Council of Canada. I am grateful to the archivists of many libraries and to *les Chevaliers des Album de Statistique Graphique*: Antoine de Falguerolles, Ruddy Ostermann, Gilles Palsky, Ian Spence, Antony Unwin, and Howard Wainer for historical information, images, and helpful suggestions.

domains, that gave rise to the many branches we see today, are often not well known, and have never been assembled in a single garden, to be studied or admired.

This chapter provides an overview of the intellectual history of data visualization from medieval to modern times, describing and illustrating some significant advances along the way. It is based on what I call the Milestones Project, an attempt to provide a broadly comprehensive and representative catalog of important developments in *all* fields related to the history of data visualization.

There are many historical accounts of developments within the fields of probability (Hald, 1990), statistics (Pearson, 1978, Porter, 1986, Stigler, 1986), astronomy (Riddell, 1980), cartography (Wallis and Robinson, 1987), which relate to, *inter alia*, some of the important developments contributing to modern data visualization. There are other, more specialized accounts, which focus on the early history of graphic recording (Hoff and Geddes, 1959, 1962), statistical graphs (Funkhouser, 1936, 1937, Royston, 1970, Tilling, 1975), fitting equations to empirical data (Farebrother, 1999), economics and time-series graphs (Klein, 1997), cartography (Friis, 1974, Kruskal, 1977) and thematic mapping (Robinson, 1982, Palsky, 1996), and so forth; Robinson (Robinson, 1982, Ch. 2) presents an excellent overview of some of the important scientific, intellectual, and technical developments of the 15th–18th centuries leading to thematic cartography and statistical thinking. Wainer and Velleman (2001) provide a recent account of some of the history of statistical graphics.

But there are no accounts that span the entire development of visual thinking and the visual representation of data, and which collate the contributions of disparate disciplines. In as much as their histories are intertwined, so too should be any telling of the development of data visualization. Another reason for interweaving these accounts is that practitioners in these fields today tend to be highly specialized, and unaware of related developments in areas outside their domain, much less their history.

## 2 Milestones Tour

*Every picture tells a story.*

—Rod Stewart, 1971

In organizing this history, it proved useful to divide history into epochs, each of which turned out to be describable by coherent themes and labels. This division is, of course somewhat artificial, but it provides the opportunity to characterize the accomplishments in each period in a general way, before describing some of them in more detail. Figure 1, discussed in Section 3.2, provides a graphic overview of the epochs I describe in the subsections below, showing the frequency of events considered milestones in the periods of this history. For now, it suffices to note the labels attached to these epochs, a steady rise from the early 18<sup>th</sup> century to the late 19<sup>th</sup> century, with a curious wiggle thereafter.

In the larger picture—recounting the history of data visualization—it turns out that many of the milestones items have a story to be told: What motivated this development? What was the communication goal? How does it relate to other developments—What were the pre-cursors? How has this idea been used or re-invented today? Each section below tries to illustrate the general themes with a few exemplars. In particular, this account attempts to tell a few representative stories of these periods, rather than to try to be comprehensive.

For reasons of economy, only a limited number of images could be printed here, and these only in black and white. Others are referred to by web links, mostly from the Milestones

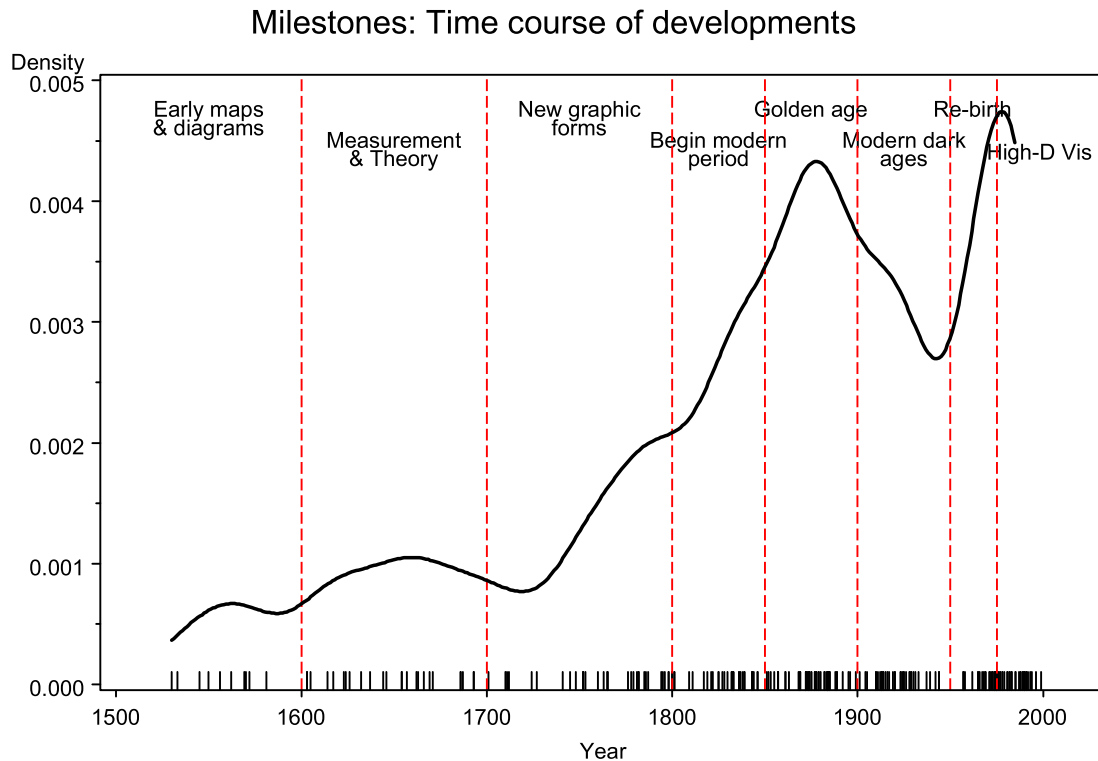


Figure 1: The time distribution of events considered milestones in the history of data visualization, shown by a rug plot and density estimate.

Project, <http://www.math.yorku.ca/SCS/Gallery/milestone/>, where a color version of this chapter will also be found.

## 2.1 Pre-17th Century: Early maps and diagrams

The earliest seeds of visualization arose in geometric diagrams, in tables of the positions of stars and other celestial bodies, and in the making of maps to aid in navigation and exploration. The idea of coordinates was used by ancient Egyptian surveyors in laying out towns, earthly and heavenly positions were located by something akin to latitude and longitude at least by 200 BC, and the map projection of a spherical earth into latitude and longitude by Claudius Ptolemy [c. 85–c. 165] in Alexandria would serve as reference standards until the 14<sup>th</sup> century.

Among the earliest graphical depictions of quantitative information is an anonymous 10<sup>th</sup> century multiple time-series graph of the changing position of the seven most prominent heavenly bodies over space and time (Figure 2), described by Funkhouser (1936) and reproduced in Tufte (1983, p. 28). The vertical axis represents the inclination of the planetary orbits, the horizontal axis shows time, divided into thirty intervals. The sinusoidal variation, with different periods is notable, as is the use of a grid, suggesting both an implicit notion of a coordinate system, and something akin to graph paper, ideas that would not be fully developed until the 1600–1700s.

In the 14<sup>th</sup> century, the idea of a plotting a theoretical function (as a proto bar graph), and

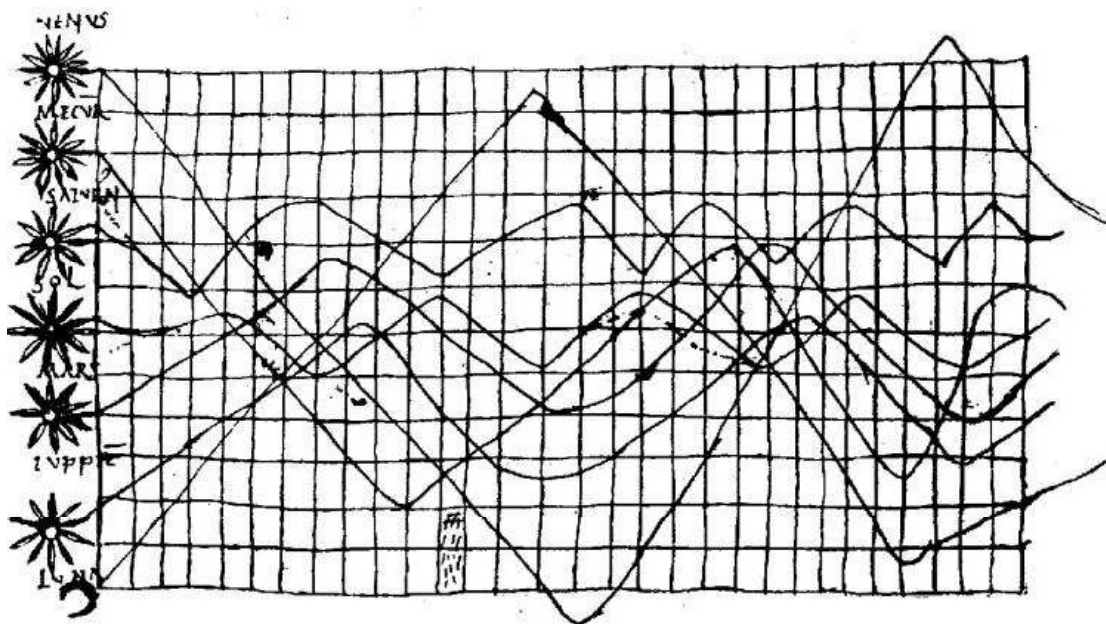


Figure 2: Planetary movements shown as cyclic inclinations over time, by an unknown astronomer, appearing in a 10<sup>th</sup> century appendix to commentaries by A. T. Macrobius on Cicero's *In Somnium Scipionus*. Source: Funkhouser (1936, p. 261).

the logical relation between tabulating values and plotting them appeared in a work by Nicole Oresme [1323–1382] Bishop of Liseus<sup>1</sup> (Oresme, 1482, 1968), followed somewhat later by the idea of a theoretical graph of distance vs. speed by Nicolas of Cusa.

By the 16<sup>th</sup> century, techniques and instruments for precise observation and measurement of physical quantities, and geographic and celestial position were well-developed (for example, a “wall quadrant” constructed by Tycho Brahe [1546–1601], covering an entire wall in his observatory). Particularly important were the development of triangulation and other methods to determine mapping locations accurately (Frisius, 1533, Tartaglia, 1556). As well, we see initial ideas for capturing images directly (the camera obscura, used by Reginer Gemma-Frisius in 1545 to record an eclipse of the sun), the recording of mathematical functions in tables (trigonometric tables by Georg Rheticus, 1550), and the first modern cartographic atlas (*Teatrum Orbis Terrarum* by Abraham Ortelius, 1570). These early steps comprise the beginnings of data visualization.

## 2.2 1600-1699: Measurement and theory

Among the most important problems of the 17<sup>th</sup> century were those concerned with physical measurement— of time, distance, and space— for astronomy, surveying, map making, navigation and territorial expansion. This century also saw great new growth in theory and the dawn of practical application— the rise of analytic geometry and coordinate systems (Descartes and

<sup>1</sup>Funkhouser (1936, p. 277) was sufficiently impressed with Oresme's grasp of the relation between functions and graphs that he remarked, “if a pioneering contemporary had collected some data and presented Oresme with actual figures to work upon, we might have had statistical graphs four hundred years before Playfair.”

Fermat), theories of errors of measurement and estimation (initial steps by Galileo in the analysis of observations on Tycho Brahe's star of 1572 (Hald, 1990, §10.3)), the birth of probability theory (Pascal and Fermat), and the beginnings of demographic statistics (John Graunt) and “political arithmetic” (William Petty)—the study of population, land, taxes, value of goods, etc. for the purpose of understanding the wealth of the state.

Early in this century, Christopher Scheiner (1630, recordings from 1611) introduced an idea Tufté (1983) would later call the principle of “small multiples” to show the changing configurations of sunspots over time, shown in Figure 3. The multiple images depict the recordings of sunspots from 23 October 1611 until 19 December of that year. The large key in the upper left identifies seven groups of sunspots by the letters A–F. These groups are similarly identified in the 37 smaller images, arrayed left-to-right and top-to-bottom below.

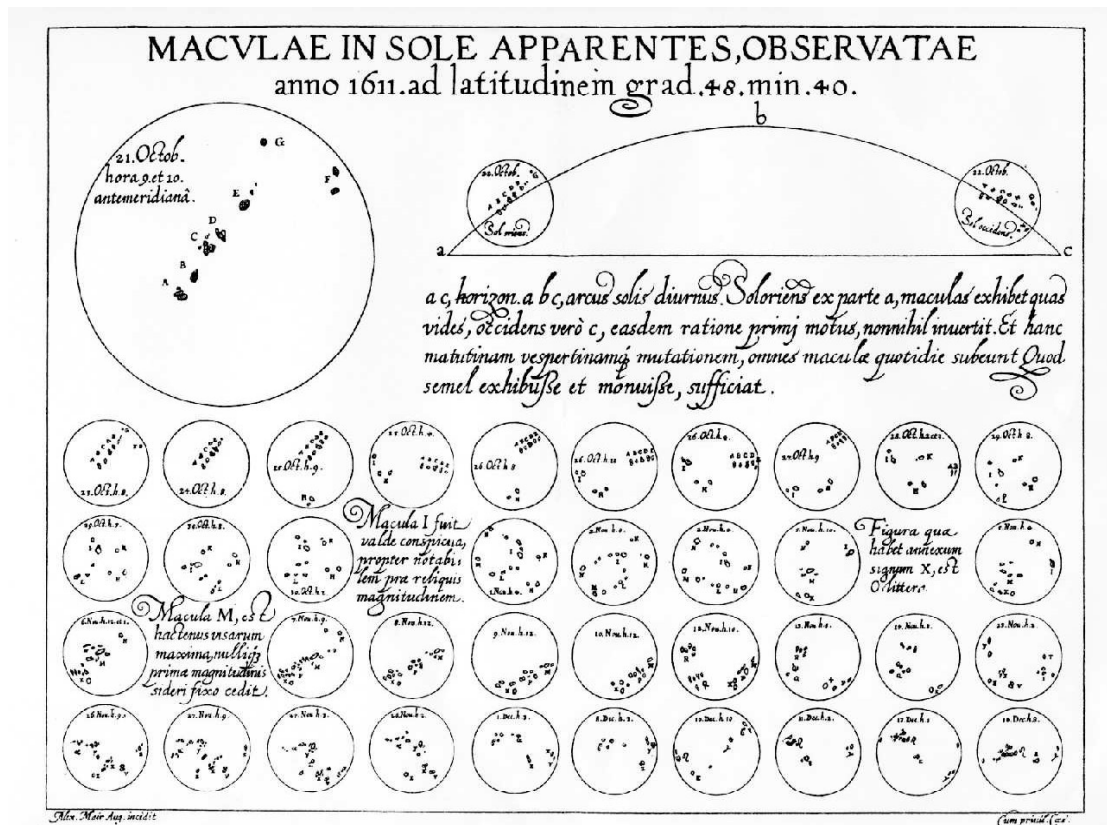


Figure 3: Scheiner's 1626 representation of the changes in sunspots over time. Source: Scheiner (1630).

Another noteworthy example (Figure 4) shows a 1644 graphic by Michael Florent van Langren [1600–1675], a Flemish astronomer to the court of Spain, believed to be the first visual representation of statistical data (Tufté, 1997, p. 15). At that time, lack of a reliable means to determine longitude at sea hindered navigation and exploration.<sup>2</sup> This 1D line graph shows all

<sup>2</sup>For navigation, latitude could be fixed from star inclinations, but longitude required accurate measurement of time at sea, an unsolved problem until 1765 with the invention of a marine chronometer by John Harrison. See Sobel (1996) for a popular account.



12 known estimates of the difference in longitude between Toledo and Rome, and the name of the astronomer (Mercator, Tycho Brahe, Ptolemy, etc.) who provided each observation.

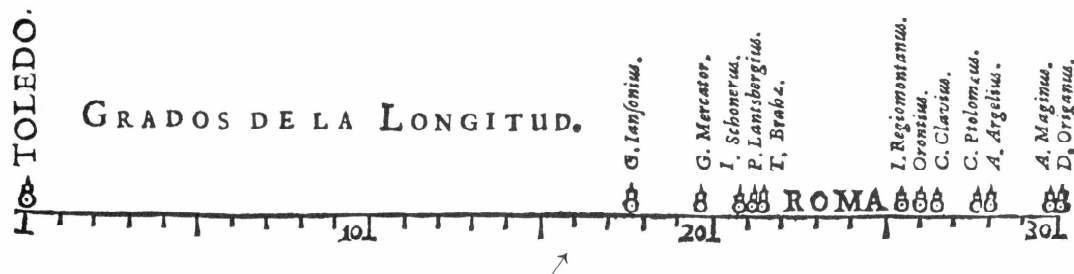


Figure 4: Langren's 1644 graph of determinations of the distance, in longitude, from Toledo to Rome. The correct distance is  $16^{\circ}30'$ . Source: Tufte (1997, p. 15).

What is notable is that van Langren could have presented this information in various tables—ordered by author to show provenance, by date to show priority, or by distance. However, only a graph shows the wide variation in the estimates; note that the range of values covers nearly half the length of the scale. Van Langren took as his overall summary the center of the range, where there happened to be a large enough gap for him to inscribe “ROMA.” Unfortunately, all of the estimates were biased upwards; the true distance ( $16^{\circ}30'$ ) is shown by the arrow. Van Langren's graph is also a milestone as the earliest-known exemplar of the principle of “effect ordering for data display” (Friendly and Kwan, 2003).

In the 1660s, the systematic collection and study of social data began in various European countries, under the rubric of “political arithmetic” (John Graunt 1662 and William Petty 1665), with the goals of informing the state about matters related to wealth, population, agricultural land, taxes and so forth,<sup>3</sup> as well as for commercial purposes such as insurance and annuities based on life tables (Jan de Witt, 1671). At approximately the same time, the initial statements of probability theory around 1654 (see Ball (1908)) together with the idea of coordinate systems were applied by Christiaan Huygens in 1669 to give the first graph of a continuous distribution function<sup>4</sup> (from Gaunt's life table based on the bills of mortality). The mid 1680s saw the first bivariate plot derived from empirical data, a theoretical curve relating barometric pressure to altitude, and the first known weather map,<sup>5</sup> showing prevailing winds on a map of the earth (Halley, 1686).

By the end of this century, the necessary elements for the development of graphical methods were at hand—some real data of significant interest, some theory to make sense of them, and a few ideas for their visual representation. Perhaps more importantly, one can see this century as giving rise to the beginnings of visual thinking, as illustrated by the examples of Scheiner and van Langren.

<sup>3</sup>For example, Graunt (1662) used his tabulations of London births and deaths from parish records and the bills of mortality to estimate the number of men the king would find available in the event of war (Klein, 1997, 43–47).

<sup>4</sup>Image: <http://math.yorku.ca/SCS/Gallery/images/huygens-graph.gif>

<sup>5</sup>Image: <http://math.yorku.ca/SCS/Gallery/images/halleyweathermap-1686.jpg>





the form of an annotated chart of all of history (from Creation) on a 54-foot scroll (Ferguson, 1991). Joseph Priestley, presumably independently, used a more convenient form to show first a timeline chart of biography (lifespans of 2,000 famous people, 1200 B.C. to 1750 A.D., Priestley (1765)), and then a detailed chart of history (Priestley, 1769).

The use of geometric figures (squares or rectangles) and cartograms to compare areas or demographic quantities by Charles de Fourcroy<sup>6</sup> (1782) and August F.W. Crome (1785) provided another novel visual encoding for quantitative data using superimposed squares to compare the areas of European states.

As well, several technological innovations provided necessary ingredients for the production and dissemination of graphic works. Some of these facilitated the reproduction of data images, such as three-color printing, invented by Jacob le Blon in 1710 and lithography by Aloys Senefelder in 1798. Of the latter, Robinson (1982, p. 57) says “the effect was as great as the introduction [of the Xerox machine].” Yet, likely due to expense, most of these new graphic forms appeared in publications with limited circulation, unlikely to attract wide attention.

A prodigious contributor to the use of the new graphical methods, Johann Lambert [1728–1777] introduced the ideas of curve fitting and interpolation from empirical data points. He used various sorts of line graphs and graphical tables to show periodic variation, for example, in air and soil temperature.<sup>7</sup>

William Playfair [1759–1823] is widely considered the inventor of most of the graphical forms widely used today—first the line graph and bar chart (Playfair, 1786), later the pie chart and circle graph (Playfair, 1801). Figure 6 shows a creative combination of different visual forms: circles, pies and lines, re-drawn from Playfair (1801, Plate 2).

The use of two separate vertical scales for different quantities (population and taxes) is today considered a sin in statistical graphics (you can easily jiggle either scale to show different things). But Playfair used this device to good effect here to try to show taxes per capita in various nations and argue that the British were overtaxed, compared with others. But alas, showing simple numbers by a graph was hard enough for Playfair—he devoted several pages of text in Playfair (1786) describing how to read and understand a line graph. The idea of calculating and graphing rates and other indirect measurements was still to come.

In this figure the left axis and line on each circle/pie graph shows population, while the right axis and line shows taxes. Playfair intended that the *slope* of the line connecting the two would depict the rate of taxation directly to the eye; but, of course, the slope also depends on the diameters of the circles. Playfair’s graphic sins can perhaps be forgiven here, because the graph clearly shows the slope of the line for Britain to be in the opposite direction of those for the other nations.

A somewhat later graph (Playfair, 1821), shown in Figure 7, exemplifies the best that Playfair had to offer with these graphic forms. Playfair used three parallel time series to show the price of wheat, weekly wages, and reigning monarch over a ~250 year span from 1565 to 1820, and used this graph to argue that workers had become better off in the most recent years.

By the end of this century (1794), the utility of graphing in scientific applications prompted a Dr. Buxton in London to patent and market printed coordinate paper; curiously, a patent for lined notepaper was not issued until 1815. The first known published graph using coordinate paper is one of periodic variation in barometric pressure (Howard, 1800). Nevertheless, graphing of

<sup>6</sup> Image: <http://math.yorku.ca/SCS/Gallery/images/palsky/defourcroy.jpg>

<sup>7</sup> Image: <http://www.journals.uchicago.edu/Isis/journal/demo/v000n000/000000/fg7.gif>

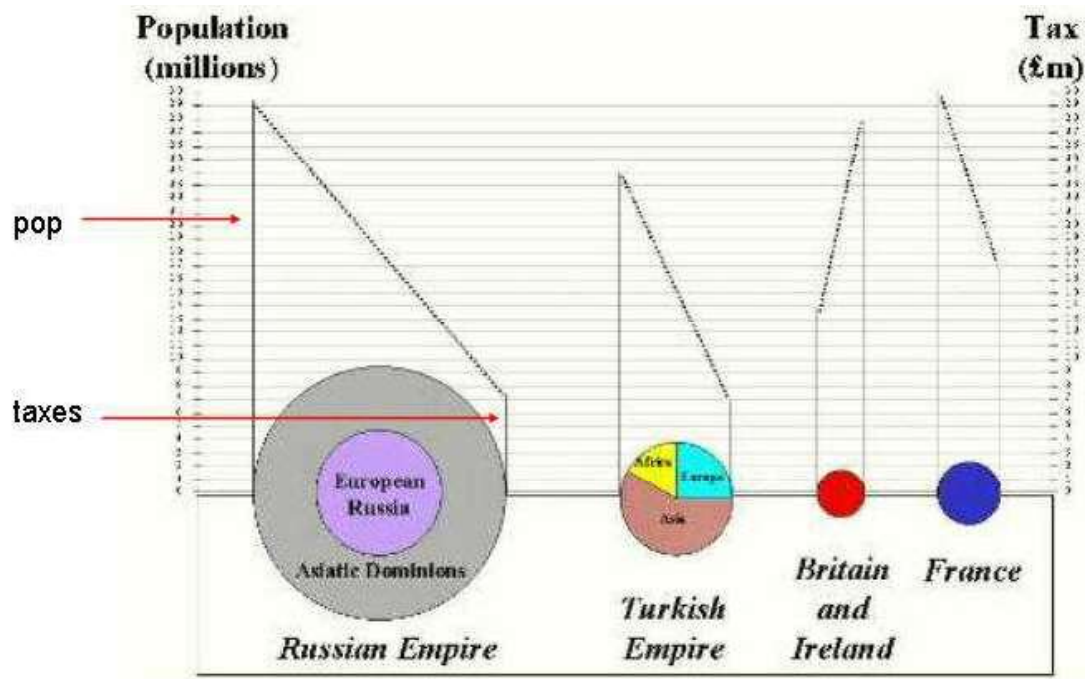


Figure 6: Re-drawn version of a portion of Playfair's 1801 pie-circle-line chart, comparing population and taxes in several nations.

data would remain rare for another 30 or so years,<sup>8</sup> perhaps largely because there wasn't much data (apart from widespread astronomical, geodetic, and physical measurement) of sufficient complexity to require new methods and applications. Official statistics, regarding population and mortality, and economic data were generally fragmentary and often not publicly available. This would soon change.

## 2.4 1800-1850: Beginnings of modern graphics

With the fertilization provided by the previous innovations of design and technique, the first half of the 19<sup>th</sup> century witnessed explosive growth in statistical graphics and thematic mapping, at a rate which would not be equalled until modern times.

In statistical graphics, all of the modern forms of data display were invented: bar and pie charts, histograms, line graphs and time-series plots, contour plots, scatterplots, and so forth. In thematic cartography, mapping progressed from single maps to comprehensive atlases, depicting data on a wide variety of topics (economic, social, moral, medical, physical, etc.), and introduced a wide range of novel forms of symbolism. During this period graphical analysis of natural and physical phenomena (lines of magnetism, weather, tides, etc.) began to appear regularly in scientific publications as well.

In 1801, the first geological maps were introduced in England by William Smith [1769–1839], setting the pattern for geological cartography or “stratigraphic geology” (Smith, 1815).

<sup>8</sup>William Herschel (1833), in a paper that describes the first instance of a modern scatterplot, devoted three pages to a description of plotting points on a grid.

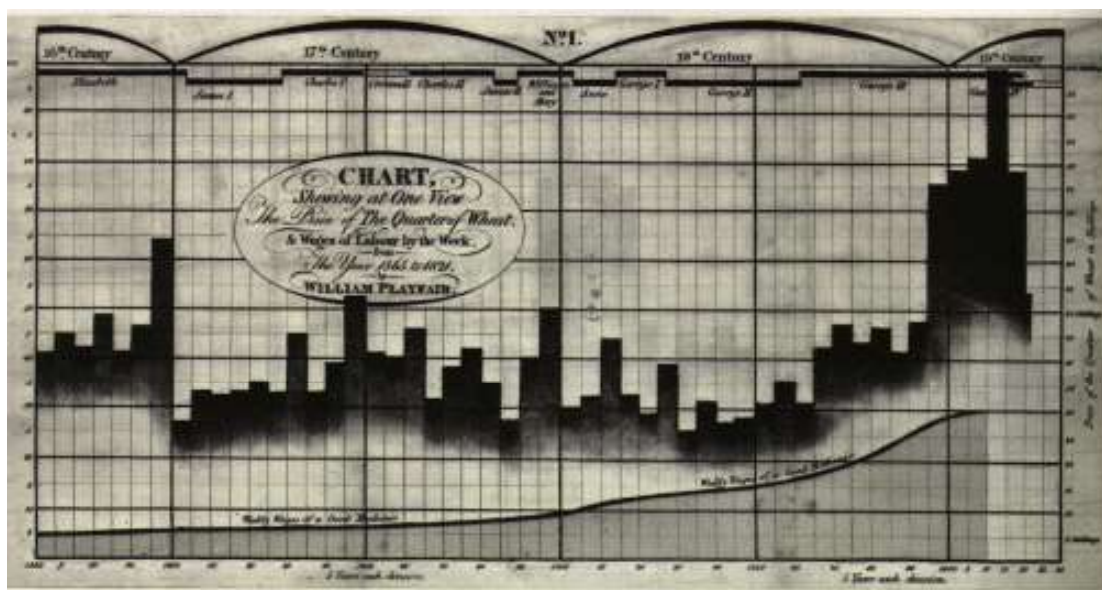


Figure 7: William Playfair's 1821 time series graph of prices, wages, and ruling monarch over a 250 year period. Source: Playfair (1821), image from Tufte (1983, p. 34)

These and other thematic maps soon led to new ways to show quantitative information on maps, and, equally importantly, to new domains for graphically-based inquiry.

In the 1820s, Baron Charles Dupin [1784–1873] invented the use of continuous shadings (from white to black) to show the distribution and degree of illiteracy in France (Dupin, 1826)—the first unclassed choropleth map,<sup>9</sup> and perhaps the first modern-style thematic statistical map (Palsky, 1996, p. 59). Later given the lovely title, “Carte de la France obscure et la France éclairée,” it attracted wide attention, and was also perhaps the first application of graphics in the social realm.

More significantly, in 1825, the Ministry of Justice in France instituted the first centralized national system of crime reporting, collected quarterly from all departments and recording the details of every charge laid before the French courts. In 1833, André-Michel Guerry, a lawyer with a penchant for numbers used this data (along with other data on literacy, suicides, donations to the poor and other “moral” variables) to produce a seminal work on the moral statistics of France (Guerry, 1833)—a work that (along with Quetelet (1831, 1835)) can be regarded as the foundation of modern social science.<sup>10</sup>

Guerry used maps in a style similar to Dupin to compare the ranking of departments on pairs of variables, notably crime vs. literacy, but other pairwise variable comparisons were made.<sup>11</sup> He used these to argue that the lack of an apparent (negative) relation between crime

<sup>9</sup> Image: <http://math.yorku.ca/SCS/Gallery/images/dupin2.gif>

<sup>10</sup> Guerry showed that rates of crime, when broken down by department, type of crime, age and gender of the accused and other variables, remained remarkably consistent from year to year, yet varied widely across departments. He used this to argue that such regularity implied the possibility of establishing social laws, much as the regularity of natural phenomena implied physical ones. Guerry also pioneered the study of suicide, with tabulations of suicides in Paris, 1827–1830, by sex, age, education, profession, etc. and a content analysis of suicide notes as to presumed motives.

<sup>11</sup> Today, one would use a scatterplot, but that graphic form was only just invented (Herschel, 1833) and would not

and literacy contradicted the arm-chair theories of some social reformers who had argued that the way to reduce crime was to increase education.<sup>12</sup> Guerry's maps and charts made somewhat of an academic sensation both in France and the rest of Europe; he later exhibited several of these at the 1851 London Exhibition, and carried out a comparative study of crime in England and France (Guerry, 1864), for which he was awarded the Moynton Prize in statistics by the French Academy of Sciences.<sup>13</sup> But Guerry's systematic and careful work was unable to shine in the shadows cast by Adolphe Quetelet, who regarded moral and social statistics as his own domain.

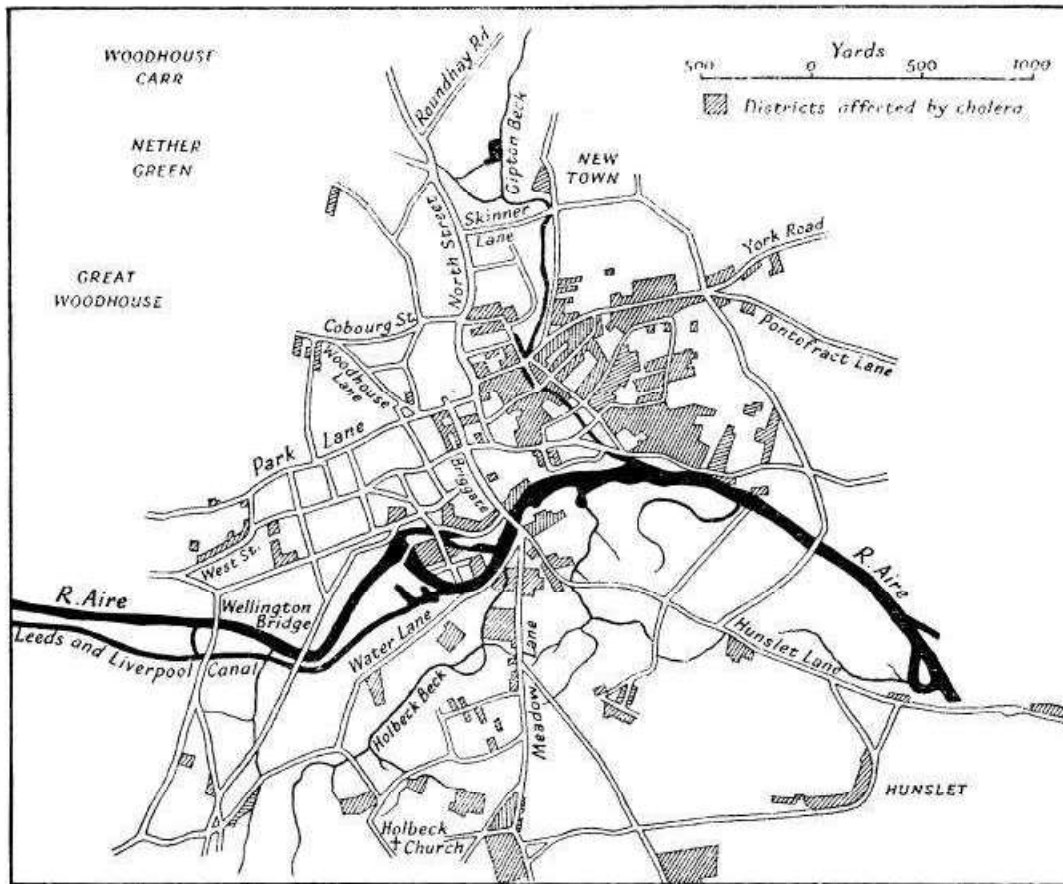


Figure 8: A portion of Dr. Robert Baker's cholera map of Leeds, 1833, showing the districts affected by cholera. *Source:* Gilbert (1958, Fig. 2).

In October 1831, the first case of asiatic cholera occurred in Great Britain, and over 52,000 enter common usage for another 50 years; see Friendly and Denis (2005).

<sup>12</sup>Guerry seemed reluctant to take sides. He also contradicted the social conservatives who argued for the need to build more prisons or impose more severe criminal sentences. See Whitt (2002).

<sup>13</sup>Among the 17 plates in this last work, seven pairs of maps for England and France each included sets of small line graphs to show trends over time, decompositions by subtype of crime and sex, distributions over months of the year, and so forth. The final plate, on general causes of crime is an incredibly detailed and complex multivariate semi-graphic display attempting to relate various types of crimes to each other, to various social and moral aspects (instruction, religion, population) as well as to their geographic distribution.



people died in the epidemic that ensued over the next 18 months or so (Gilbert, 1958). Subsequent cholera epidemics in 1848–1849 and 1853–1854 produced similarly large death tolls, but the water-borne cause of the disease was unknown until 1855 when Dr. John Snow produced his famous dot map<sup>14</sup> (Snow, 1855) showing deaths due to cholera clustered around the Broad Street pump in London. This was indeed a landmark graphic discovery, but it occurred at the end of the period, roughly 1835–1855, that marks a high-point in the application of thematic cartography to human (social, medical, ethnic) topics. The first known disease map of cholera (Figure 8), due to Dr. Robert Baker (1833), shows the districts of Leeds “affected by cholera” in the particularly severe 1832 outbreak.

I show this figure to make another point— why Baker’s map did not lead to a “eureka” experience, while John Snow’s did. Baker used a town plan of Leeds that had been divided into districts. Of a population of 76,000 in all of Leeds, Baker mapped the 1800 cholera cases by hatching in red “the districts in which the cholera had prevailed.” In his report, he noted an association between the disease and living conditions: “how exceedingly the disease has prevailed in those parts of the town where there is a deficiency, often an entire want of sewage, drainage, and paving” (Baker, 1833, p. 10). Baker did not indicate the incidence of disease on his map, nor was he equipped to display *rates* of disease (in relation to population density)<sup>15</sup> and his knowledge of possible causes, while definitely on the right track, was both weak and implicit (not analyzed graphically or by other means). It is likely that some, perhaps tenuous, causal indicants or evidence were available to Baker, but he was unable to connect the dots, or see a geographically distributed outcome in relation to geographic factors in even the simple ways that Guerry had tried.

At about the same time, ~1830–1850, the use of graphs began to become recognized in some official circles for economic and state planning— where to build railroads and canals? what is the distribution of imports and exports? This use of graphical methods is no better illustrated than in the works of Charles Joseph Minard [1781–1870], whose prodigious graphical inventions led Funkhouser (1937) to call him the Playfair of France. To illustrate, we choose (with some difficulty) an 1844 “tableau-graphique” (Figure 9) by Minard, an early progenitor of the modern mosaic plot (Friendly, 1994). On the surface, mosaic plots descend from bar charts, but Minard introduced two simultaneous innovations: the use of divided and proportional-width bars so that area had a concrete visual interpretation. The graph shows the transportation of commercial goods along one canal route in France by variable-width, divided bars (Minard, 1844).

In this display the width of each vertical bar shows distance along this route; the divided bar segments have height  $\sim$  amount of goods of various types (shown by shading), so the area of each rectangular segment is proportional to cost of transport. Minard, a true visual engineer (Friendly, 2000), developed such diagrams to argue visually for setting differential price rates for partial vs. complete runs. Playfair had tried to make data ‘speak to the eyes,’ but Minard wished to make them ‘calculer par l’œil’ as well.

It is no accident that, in England, outside the numerous applications of graphical methods in the sciences, there was little interest in or use of graphs among statisticians (or “statists” as

<sup>14</sup> Image: <http://www.math.yorku.ca/SCS/Gallery/images/snow4.jpg>

<sup>15</sup> The German geographer Augustus Petermann produced a “Cholera map of the British Isles” in 1852 using national data from the 1831–1832 epidemic, (image: <http://images.rgs.org/webimages/0/0/10000/1000/800/S0011888.jpg>) shaded in proportion to the relative rate of mortality using class intervals ( $< 1/35, 1/35 : 1/100, 1/100 : 1/200, \dots$ ). No previous disease map allowed determination of the range of mortality in any given area.

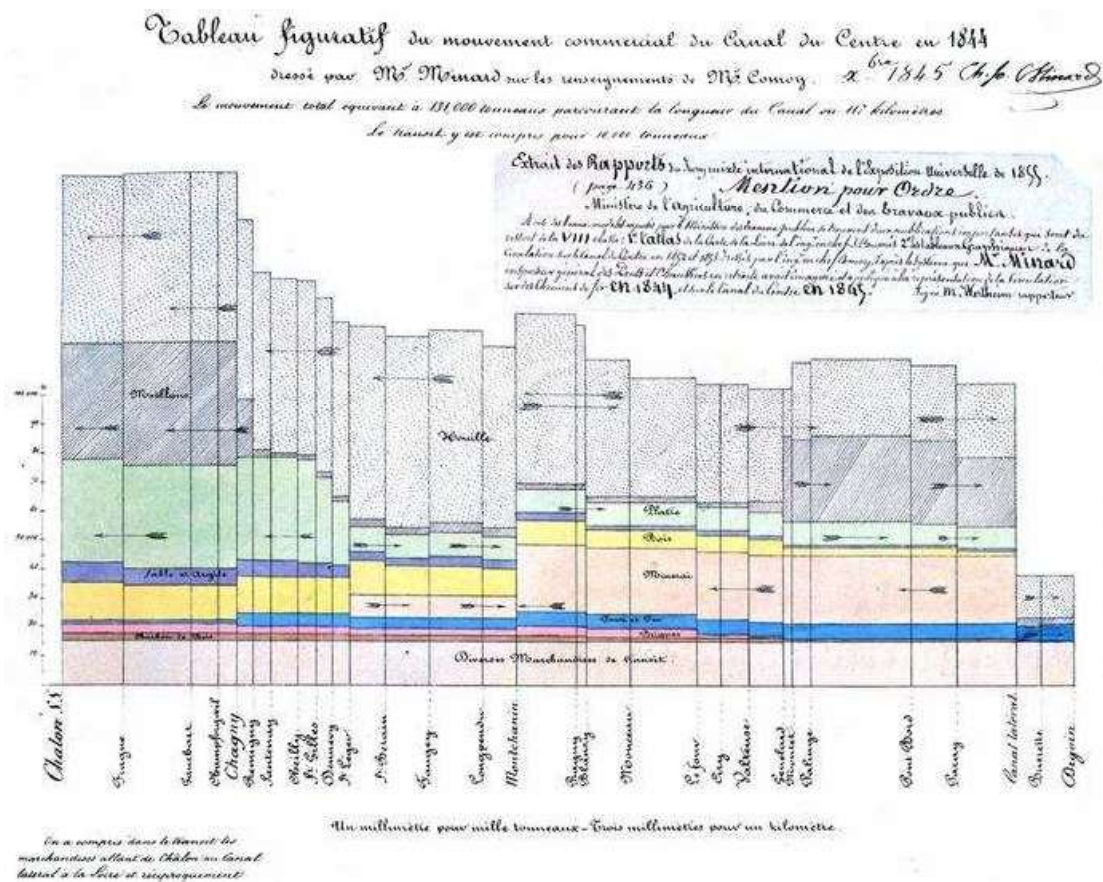


Figure 9: Minard's Tableau Graphique, showing the transportation of commercial goods along the Canal du Centre (Chalon–Dijon). Intermediate stops are spaced by distance, and each bar is divided by type of goods, so the area of each tile represents the cost of transport. Arrows show the direction of transport. *Source:* ENPC:5860/C351 (Col. et cliché ENPC; used by permission).

they called themselves). If there is a continuum ranging from “graph people” to “table people,” British statisticians and economists were philosophically more table-inclined, and looked upon graphs with suspicion up to the time of William Stanley Jevons around 1870 ([Maas and Morgan, 2005](#)). Statistics should be concerned with the recording of “facts relating to communities of men which are capable of being expressed by numbers” ([Mouat, 1885](#), p.15), leaving the generalization to laws and theories to others. Indeed, this view was made abundantly clear in the logo of the Statistical Society of London (now the Royal Statistical Society): a banded sheaf of wheat, with the motto *Aliis Exterendum*—to others to flail the wheat. Making graphs, it seemed, was too much like bread-making.

## 2.5 1850–1900: The Golden Age of statistical graphics

By the mid-1800s, all the conditions for the rapid growth of visualization had been established—a “perfect storm” for data graphics. Official state statistical offices were established throughout

Europe, in recognition of the growing importance of numerical information for social planning, industrialization, commerce, and transportation. Statistical theory, initiated by Gauss and Laplace, and extended to the social realm by Guerry and Quetelet, provided the means to make sense of large bodies of data.

What started as the *Age of Enthusiasm* (Funkhouser, 1937, Palsky, 1996) for graphics ended with what can be called the *Golden Age*, with unparalleled beauty and many innovations in graphics and thematic cartography. So varied were these developments, that it is difficult to be comprehensive, but a few themes stand out.

### 2.5.1 Escaping flatland

Although some attempts to display more than two variables simultaneously had occurred earlier in multiple time-series (Playfair, 1801, Minard, 1826), contour graphs (Vauthier, 1874) and a variety of thematic maps, (e.g., Berghaus (1838)) a number of significant developments extended graphics beyond the confines of a flat piece of paper. Gustav Zeuner [1828–1907] in Germany (Zeuner, 1869), and later Luigi Perozzo [?–1875] in Italy (Perozzo, 1880) constructed 3D surface plots of population data.<sup>16</sup> The former was an axonometric projection showing various slices, while the latter (a 3D graph of population in Sweden from 1750–1875 by year and age group) was printed in red and black and designed as a stereogram.<sup>17</sup>

Contour diagrams, showing iso-level curves of 3D surfaces, had also been used earlier in mapping contexts (Nautonier, 1604, Halley, 1701, von Humboldt, 1817), but the range of problems and data to which they were applied expanded considerably over this time in attempts to understand relations among more than two data-based variables, or where the relationships are statistical, rather than functional or measured with little error. It is more convenient to describe these under Galton, below. By 1884, the idea of visual and imaginary worlds of varying number of dimensions found popular expression in Edwin Abbott's (1884) *Flatland*, implicitly suggesting possible views in four and more dimensions.

### 2.5.2 Graphical innovations

With the usefulness of graphical displays for understanding complex data and phenomena established, many new graphical forms were invented and extended to new areas of inquiry, particularly in the social realm.

Minard (1861) developed the use of divided circle diagrams on maps (showing both a total, by area, and sub-totals, by sectors, with circles for each geographic region on the map). Later he developed to an art form the use of flow lines on maps of width proportional to quantities (people, goods, imports, exports) to show movement and transport geographically. Near the end of his life, the flow map would be taken to its highest level in his famous depiction of the fate of the armies of Napoleon and Hannibal, in what Tufte (1983) would call the “best graphic ever produced.” See Friendly (2002) for a wider appreciation of Minard's work.

The social and political uses of graphics is also evidenced in the polar area charts (called “rose diagrams” or “coxcombs”) invented by Florence Nightingale [1820–1910] to wage a

<sup>16</sup> Image: <http://math.yorku.ca/SCS/Gallery/images/stereo2.jpg>

<sup>17</sup> Zeuner used one axis to show year of birth and another to show present age, with number of surviving persons on the third, vertical axis, giving a 3D surface. One set of curves thus showed the distribution of population for a given generation; the orthogonal set of curves showed the distributions across generations at a given point in time, e.g., at a census.



campaign for improved sanitary conditions in battlefield treatment of soldiers (Nightingale, 1857). They left no doubt that many more soldiers died from disease and the consequences of wounds than at the hands of the enemy. From around the same time, Dr. John Snow [1813–1858] is remembered for his use of a dot map of deaths from cholera in an 1854 outbreak in London the cholera deaths in London. Plotting the residence of each deceased provided the insight for his conclusion that the source of the outbreak could be localized to contaminated water from a pump on Broad Street, the founding innovation for modern epidemiological mapping.

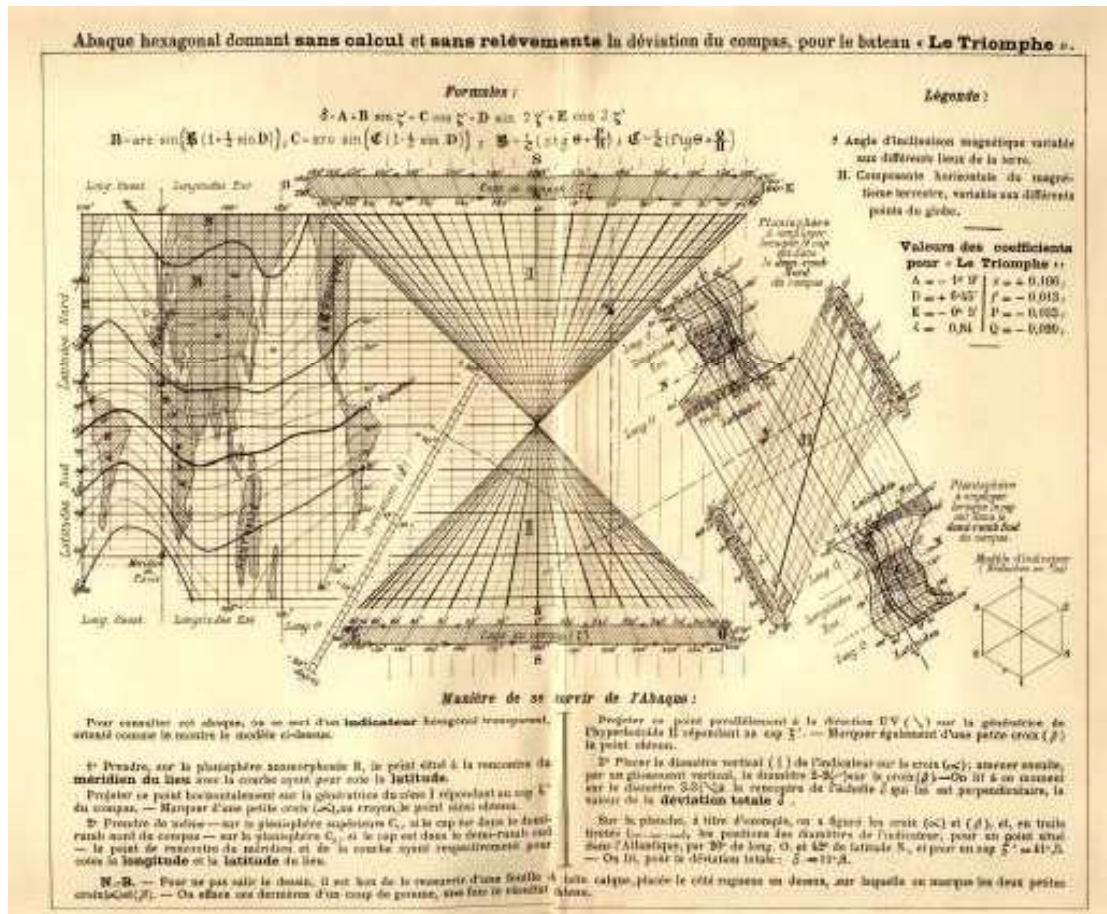


Figure 10: Lallemand's *L'abaque du bateau Le Triomphe*, allowing determination of magnetic deviation at sea without calculation. *Source:* courtesy Mme. Marie-Noëlle Maisonneuve, Les fonds anciens de la bibliothèque de l'Ecole des Mines de Paris.

Scales and shapes for graphs and maps were also transformed for a variety of purposes, leading to semi-logarithmic graphs (Jevons, 1863, 1958) to show percentage change in commodities over time, log-log plots to show multiplicative relations, anamorphic maps by Émile Cheysson (Palsky, 1996, Fig. 63-64) using deformations of spatial size to show a quantitative variable (e.g., the decrease in time to travel from Paris to various places in France over 200 years), and alignment diagrams or nomograms using sets of parallel axes. We illustrate this slice of the golden age with Figure 10, a tour-de-force graphic for determination of magnetic deviation at sea in relation to latitude and longitude without calculation (“L’Abaque Triomphe”) by

Charles Lallemand (1885), director general of the geodetic measurement of altitudes throughout France, that combines many variables into a multi-function nomogram, using 3D, juxtaposition of anamorphic maps, parallel coordinates and hexagonal grids.

### 2.5.3 Galton's contributions

Special note should be made of the varied contributions of Francis Galton [1822-1911] to data visualization and statistical graphics. Galton's role in the development of the ideas of correlation and regression are well-known. Less well-known is the role that visualization and graphing played in his contributions and discoveries.

Galton's statistical insight (Galton, 1886)— that, in a bivariate (normal) distribution (say, height of child against height of parent), (a) the isolines of equal frequency would appear as concentric ellipses, and (b) that the locus of the (regression) lines of means of  $y | x$  and of  $x | y$  were the conjugate diameters of these ellipses — was based largely on visual analysis from the application of smoothing to his data. Karl Pearson would later say, “that Galton should have evolved all this from his observations is to my mind one of the most noteworthy scientific discoveries arising from pure analysis of observations.” (Pearson, 1920, p. 37). This was only one of Galton's discoveries based on graphical methods.

In earlier work, Galton had made wide use of isolines, contour diagrams and smoothing in a variety of areas. An 1872 paper showed the use of “isodic curves” to portray the joint effects of wind and current on the distance ships at sea could travel in any direction. An 1881 “isochronic chart” (Galton, 1881) showed the time it took to reach any destination in the world from London by means of colored regions on a world map. Still later, he analyzed rates of fertility in marriages in relation to the ages of father and mother using “isogens,” curves of equal percentage of families having a child (Galton, 1894).

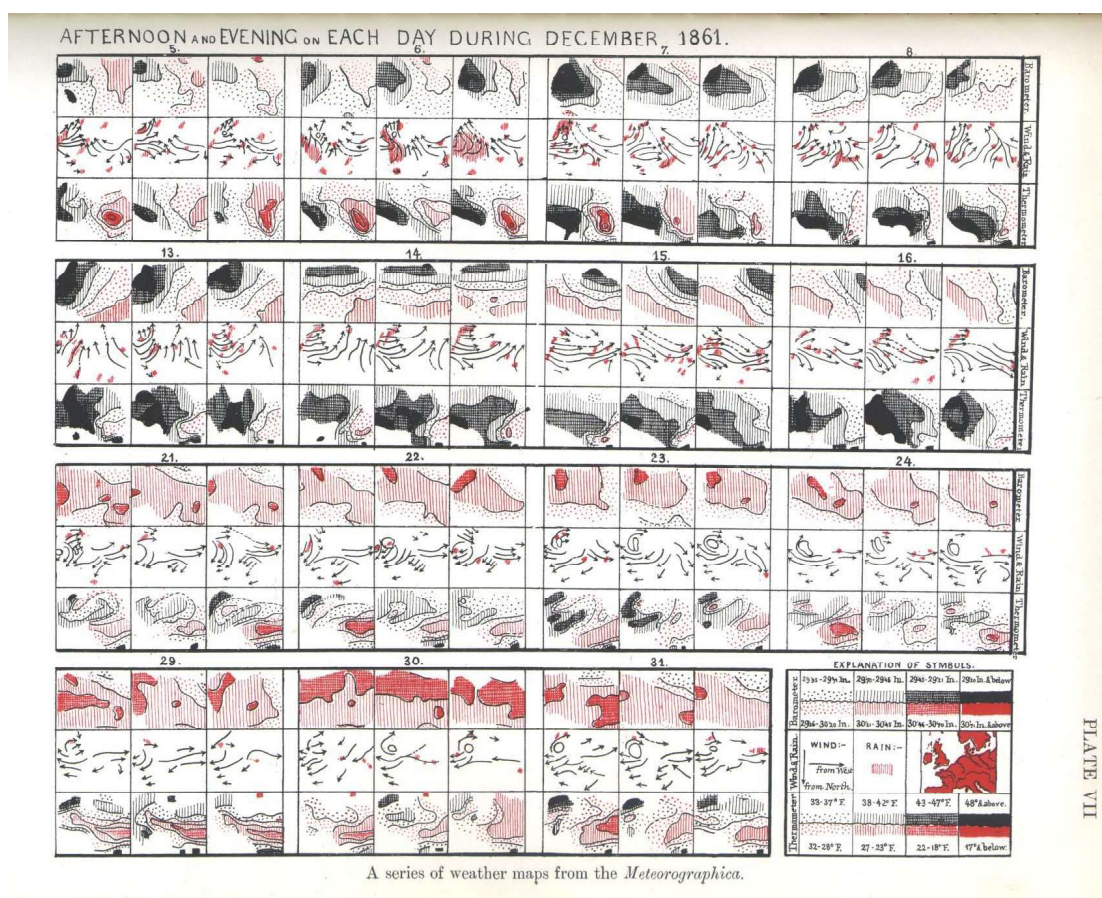
But perhaps the most notable non-statistical graphical discovery was that of the “anti-cyclonic” (counter-clockwise) pattern of winds around low-pressure regions, combined with clockwise rotations around high-pressure zones. Galton's work on weather patterns began in 1861 and was summarized in *Meteorographica* (1863). It contained a variety of ingenious graphs and maps (over 600 illustrations in total) one of which is shown in Figure 11. This remarkable chart, one of a two-page trellis-style display, shows observations on barometric pressure, wind direction, rain and temperature from 15 days in December 1861.<sup>18</sup> For each day, the  $3 \times 3$  grid shows schematic maps of Europe, mapping pressure (row 1), wind and rain (row 2) and temperature (row 3), in the morning, afternoon and evening (columns). One can clearly see the series of black areas (low pressure) on the barometric charts for about the first half of the month, corresponding to the counter-clockwise arrows in the wind charts, followed by a shift to red areas (high pressure) and more clockwise arrows. Wainer (2005, p. 56) remarks, “Galton did for the collectors of weather data what Kepler did for Tycho Brahe. This is no small accomplishment.”

### 2.5.4 Statistical Atlases

The collection, organization and dissemination of official government statistics on population, trade and commerce, social, moral and political issues became widespread in most of the coun-

<sup>18</sup> In July 1861, Galton distributed a circular to meteorologists throughout Europe, asking them to record these data synchronously, three times a day for the entire month of December, 1861. About 50 weather stations supplied the data; see Pearson (1930, p. 37–39).





(about 11 x 17 in.), and many of the plates folded out to four- or six-times that size, all printed in color and with great attention to layout and composition. We concur with Funkhouser (1937, p.336) that “the *Albums* present the finest specimens of French graphic work in the century and considerable pride was taken in them by the French people, statisticians and laymen alike.”

The subject matter of the albums largely concerned economic and financial data related to the planning, development and administration of public works— transport of passengers and freight, by rail, on inland waterways and through seaports, but also included such topics as revenues in the major theaters of Paris, attendance at the universal expositions of 1867, 1878 and 1889, changes in populations of French departments over time, and so forth.

More significantly for this account the *Albums* can also be viewed as an exquisite sampler of all the graphical methods known at the time, with significant adaptations to the problem at hand. The majority of these graphs used and extended the flow map pioneered by Minard. Others used polar forms— variants of pie and circle diagrams, star plots and rose diagrams, often overlaid on a map and extended to show additional variables of interest. Still others used sub-divided squares in the manner of modern mosaic displays (Friendly, 1994) to show the breakdown of a total (passengers, freight) by several variables. It should be noted that in almost all cases the graphical representation of the data was accompanied by numerical annotations or tables, providing precise numerical values.

The *Albums* are discussed extensively by Palsky (1996), who includes seven representative illustrations. It is hard to choose a single image here, but my favorites are surely the recursive, multi-mosaic of rail transportation for the 1884–1886 volumes, the first of which is shown in Figure 12. This cartogram uses one large mosaic (in the lower left) to show the numbers of passengers and tons of freight shipped from Paris from the four principal train stations. Of the total leaving Paris, the amounts going to each main city are shown by smaller mosaics, colored according to railway lines; of those amounts, the distribution to smaller cities is similarly shown, connected by lines along the rail routes.

Among the many other national statistical albums and atlases, those from the U.S. Census bureau also deserve special mention. The *Statistical Atlas of the Ninth Census*, produced in 1872–1874 under the direction of Francis A. Walker [1840–1897] contained 60 plates, including several novel graphic forms. The ambitious goal was to present a graphic portrait of the nation, and covered a wide range of physical and human topics: geology, minerals, weather; population by ethnic origin, wealth, illiteracy, school attendance and religious affiliation; death rates by age, sex, race and cause, prevalence of blindness, deaf mutism and insanity, and so forth. “Age pyramids” (back-to-back, bilateral frequency histograms and polygons) were used effectively to compare age distributions of the population for two classes (gender, married/single, etc.). Sub-divided squares and area-proportional pies of various forms were also used to provide comparisons among the states on multiple dimensions simultaneously (employed/unemployed, sex, schooling, occupational categories). The desire to provide for easy comparisons among states and other categorizations was expressed by arranging multiple sub-figures as “small multiples” in many plates.

Following each subsequent decennial census for 1880 to 1900, reports and statistical atlases were produced with more numerous and varied graphic illustrations. The 1898 volume from the Eleventh Census (1890), under the direction of Henry Gannett [1846–1914] contained over 400 graphs, cartograms and statistical diagrams. There were several ranked parallel coordinate

---

the International Statistical Congresses from 1872 on. He was trained as an engineer at the ENPC, and later became a professor of political economy at the École des Mines.



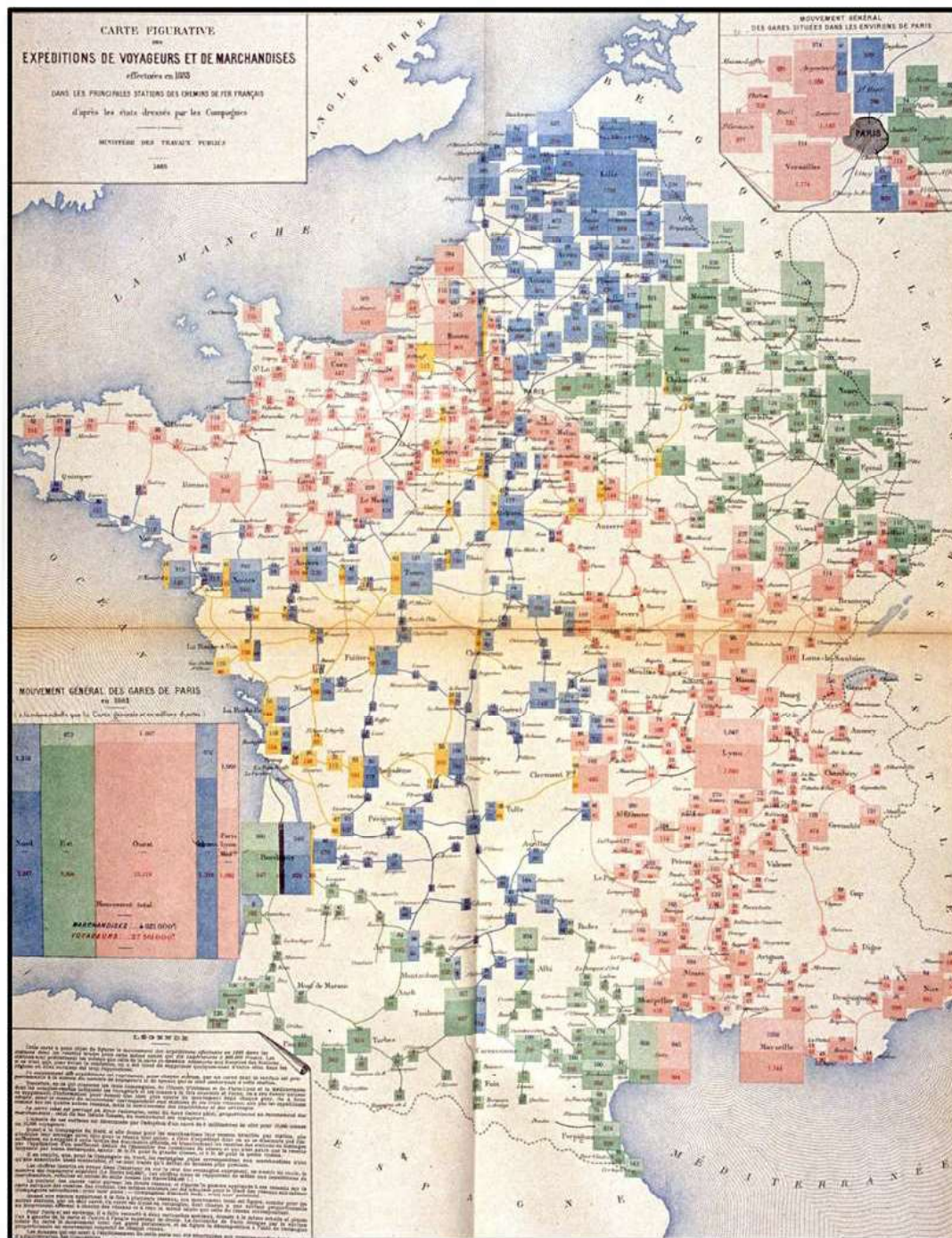


Figure 12: *Mouvement des voyageurs et des marchandises dans les principales stations de chemins de fer en 1882*. Scale: 2mm<sup>2</sup> = 10,000 passengers or tons of freight. *Source*: Album, 1884, Plate 11 (author's collection).

plots comparing states and cities over all censuses from 1790–1890. Trellis-like collections of shaded maps showed interstate migration, distributions of religious membership, deaths by known causes, and so forth.

The 1880 and 1890 volumes produced under Gannett's direction are also notable for (a) the multi-modal combination of different graphic forms (maps, tables, bar charts, bilateral polygons) in numerous plates, and (b) the consistent use of effect-order sorting ([Friendly and Kwan, 2003](#)) to arrange states or other categories in relation to what was to be shown, rather than for lookup (e.g., Alabama–Wyoming).

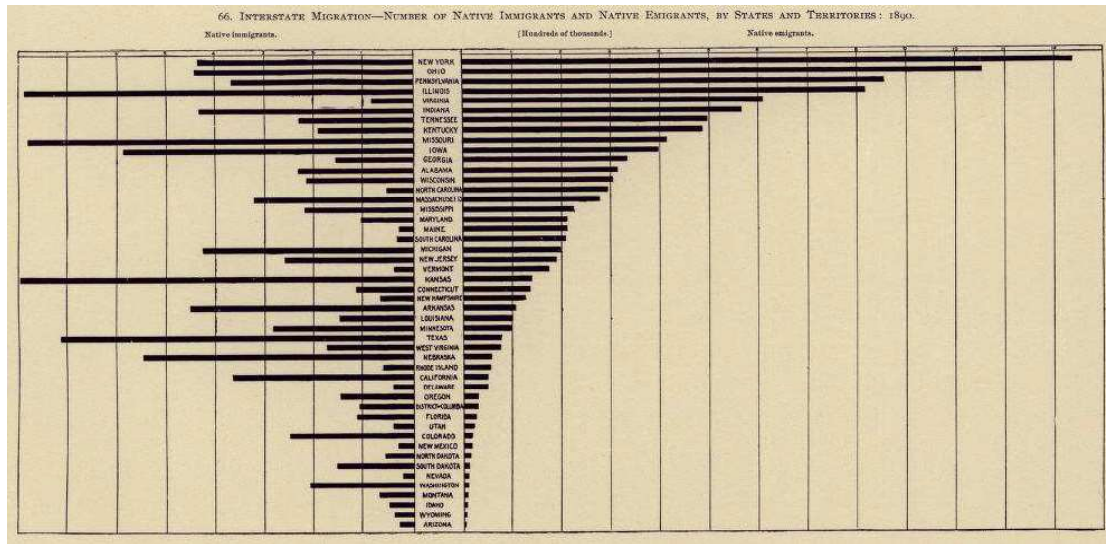


Figure 13: Interstate migration shown by back-to-back bar charts, sorted by emigration. *Source:* Statistical Atlas of the Eleventh Census, 1890, diagram 66, p. 23 (author's collection).

For example, Figure 13 shows interstate immigration in relation to emigration for the 49 states and territories in 1890. The right side shows population loss sorted by emigration, ranging from NY, Ohio, Penn. and Illinois at the top to Idaho, Wyoming and Arizona at the bottom. The left side shows where the emigrants went: Illinois, Missouri, Kansas and Texas had the biggest gains, Virginia the biggest net loss. It is clear that people were leaving the eastern states and were attracted to those of the midwest Mississippi valley. Other plates showed this data in map-based formats.

However, the Age of Enthusiasm and the Golden Age were drawing to a close. The French *Albums de Statistique Graphique* were discontinued in 1897 due to the high cost of production; statistical atlases appeared in Switzerland in 1897 and 1914, but never again. The final two U.S. Census atlases, issued after the 1910 and 1920 censuses, “were both routinized productions, largely devoid of color and graphic imagination” ([Dahmann, 2001](#)).

## 2.6 1900-1950: The modern dark ages

If the late 1800s were the “golden age” of statistical graphics and thematic cartography, the early 1900s can be called the “modern dark ages” of visualization ([Friendly and Denis, 2000](#)).

There were few graphical innovations, and, by the mid-1930s, the enthusiasm for visualization which characterized the late 1800s had been supplanted by the rise of quantification and



formal, often statistical, models in the social sciences. Numbers, parameter estimates, and, especially, those with standard errors were precise. Pictures were— well, just pictures: pretty or evocative, perhaps, but incapable of stating a “fact” to three or more decimals. Or so it seemed to many statisticians.

But it is equally fair to view this as a time of necessary dormancy, application, and popularization, rather than one of innovation. In this period statistical graphics became main stream. Graphical methods entered English<sup>20</sup> textbooks (Bowley, 1901, Peddle, 1910, Haskell, 1919, Karsten, 1925), the curriculum (Costelloe, 1915, Warne, 1916), and standard use in government (Ayres, 1919), commerce (Gantt charts and Shewart’s control charts) and science.

These textbooks contained rather detailed descriptions of the graphic method, with an appreciative and often modern flavor. For example, Sir Arthur Bowley’s (1901) *Elements of Statistics* devoted two chapters to graphs and diagrams, and discussed frequency and cumulative frequency curves (with graphical methods for finding the median and quartiles), effects of choice of scales and baselines on visual estimation of differences and ratios, smoothing of time-series graphs, rectangle diagrams in which three variables could be shown by height, width and area of bars, and “historical diagrams” in which two or more time series could be shown on a single chart for comparative views of their histories.

Bowley’s (1901, p. 151-154) example of smoothing (see Figure 14) illustrates the character of his approach. Here he plotted the total value of exports from Britain and Ireland over 1855–1899. At issue was whether exports had become stationary in the most recent years and the conclusion by Sir Robert Giffen (1899), based solely on tables of averages for successive five year periods,<sup>21</sup> that “the only sign of stationariness is an increase at a less rate in the last periods than in the earlier periods” (p. 152). To answer this, he graphed the raw data, together with curves of the moving average over three, five and ten year periods. The three- and five-year moving averages show strong evidence of an approximately 10 year cycle, and he noted, “no argument can stand which does not take account of the cycle of trade, which is not eliminated until we take decennial averages” (p. 153). To this end, he took averages of successive 10-year periods starting 1859 and drew a freehand curve “keeping as close [to the points] as possible, without making sudden changes in curvature,” giving the thick curve in Figure 14.<sup>22</sup> Support for Sir Robert’s conclusion and the evidence for a 10-year cycle owe much to this graphical treatment.

Moreover, perhaps for the first time, graphical methods proved crucial in a number of new insights, discoveries, and theories in astronomy, physics, biology, and other sciences. Among these, one may refer to (a) E. W. Maunder’s (1904) “butterfly diagram” to study the variation of sunspots over time, leading to the discovery that they were markedly reduced in frequency from 1645–1715; (b) the Hertzsprung-Russell diagram (Hertzsprung, 1911, Spence and Garrison, 1993), a log-log plot of luminosity as a function of temperature for stars, used to explain the changes as a star evolves and laying the groundwork for modern stellar physics; (c) the dis-

<sup>20</sup>The first systematic attempt to survey, describe, and illustrate available graphic methods for experimental data was Étienne Jules Marey’s (1878) *La Méthode Graphique*. Marey [1830–1904] also invented several devices for visual recording, including the sphymograph and chronophotography to record motion of birds in flight, people running, and so forth.

<sup>21</sup>Giffen, an early editor of *The Statist*, also wrote a statistical text published posthumously in 1913; it contained an entire chapter on constructing tables, but not a single graph (Klein, 1997, p. 17).

<sup>22</sup>A reanalysis of the data using a loess smoother shows that this is in fact over-smoothed, and corresponds closely to a loess window width of  $f = 0.50$ . The optimal smoothing parameter, minimizing  $AIC_C$  is  $f = 0.16$ , giving a smooth more like Bowley’s three- and five-year moving averages.



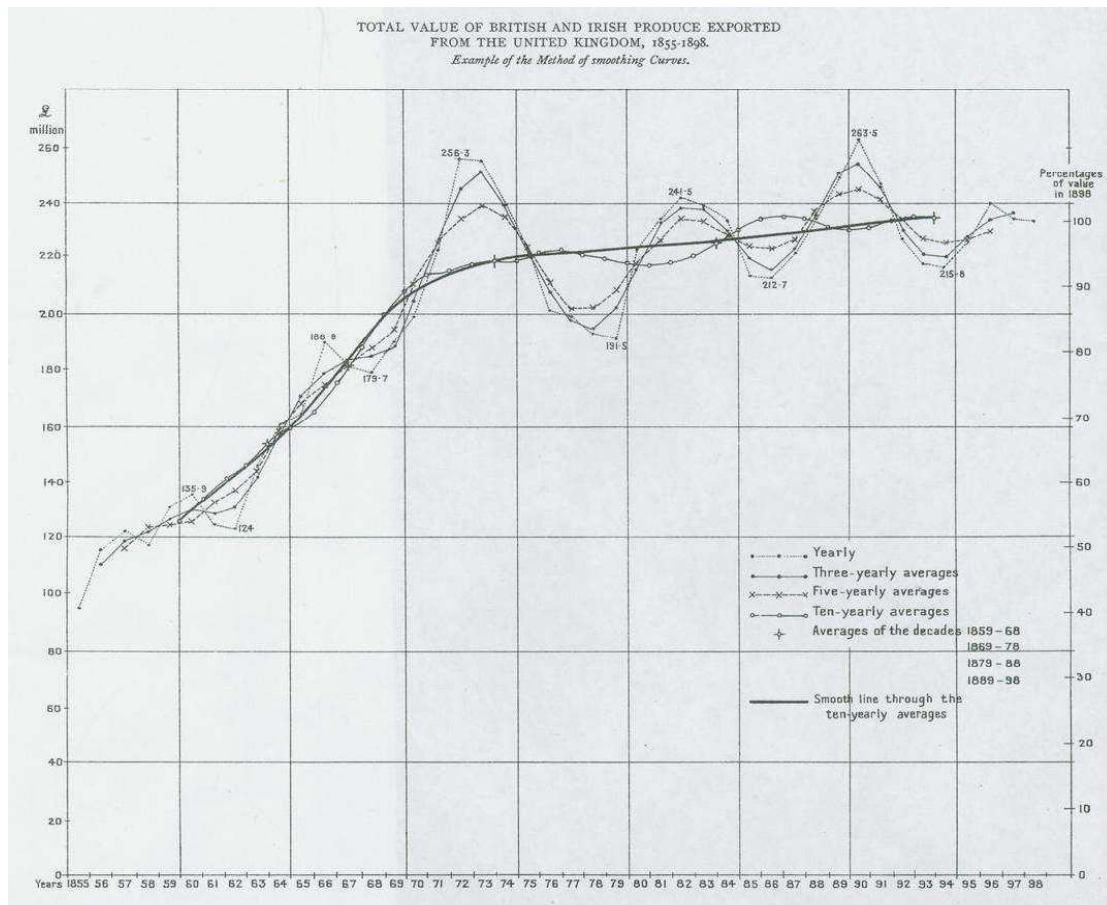


Figure 14: Arthur Bowley's demonstration of methods of smoothing a time series graph. Moving averages of three, five and ten years are compared with a freehand curve drawn through four points representing the averages of successive ten year periods. *Source: Bowley (1901, opposite p. 151).*

covery of the concept of atomic number by Henry Moseley (1913) based largely on graphical analysis. See (Friendly and Denis, 2005) for more detailed discussion of these uses.

As well, experimental comparisons of the efficacy of various graphics forms were begun (Eells, 1926, von Huhn, 1927, Washburne, 1927), a set of standards and rules for graphic presentation was finally adopted by a joint committee (Joint Committee on Standards for Graphic Presentation, 1914), and a number of practical aids to graphing were developed. In the latter part of this period, new ideas and methods for multi-dimensional data in statistics and psychology would provide the impetus to look beyond the 2D plane.

Graphic innovation was also awaiting new ideas and technology: the development of the machinery of modern statistical methodology, and the advent of the computational power and display devices which would support the next wave of developments in data visualization.

## 2.7 1950–1975: Re-birth of data visualization

Still under the influence of the formal and numerical zeitgeist from the mid-1930s on, data visualization began to rise from dormancy in the mid 1960s. This was spurred largely by three significant developments:

- In the USA, John W. Tukey [1915–2000], in a landmark paper, *The Future of Data Analysis* (Tukey, 1962), issued a call for the recognition of data analysis as a legitimate branch of statistics distinct from mathematical statistics; shortly, he began the invention of a wide variety of new, simple, and effective graphic displays, under the rubric of “Exploratory Data Analysis” (EDA)— stem-leaf plots, boxplots, hanging rootograms, two-way table displays, and so forth, many of which entered the statistical vocabulary and software implementation. Tukey’s stature as a statistician and the scope of his informal, robust, and graphical approach to data analysis were as influential as his graphical innovations. Although not published until 1977, chapters from Tukey’s EDA book (Tukey, 1977) were widely circulated as they began to appear in 1970–1972, and began to make graphical data analysis both interesting and respectable again.
- In France, Jacques Bertin [1918–] published the monumental *Semiologie Graphique* (Bertin, 1967). To some, this appeared to do for graphics what Mendeleev had done for the organization of the chemical elements, that is, to organize the visual and perceptual elements of graphics according to the features and relations in data. In a parallel but separate steam, an exploratory and graphical approach to multidimensional data (“L’analyse des données”) begun by Jean-Paul Benzécri [1932–] provided French and other European statisticians with an alternative, visually-based view of what statistics was about.
- But the skills of hand-drawn maps and graphics had withered during the dormant “modern dark ages” of graphics (though nearly every figure in Tukey’s *EDA* (Tukey, 1977) was, by intention, hand-drawn). Computer processing of statistical data began in 1957 with the creation of FORTRAN, the first high-level language for computing. By the late 1960s, widespread mainframe university computers offered the possibility to construct old and new graphic forms by computer programs. Interactive statistical applications, e.g., Fowlkes (1969), Fishkeller *et al.* (1974) and true high-resolution graphics were developed, but would take a while to enter common use.

By the end of this period significant intersections and collaborations would begin: (a) computer science research (software tools, C language, UNIX, etc.) at Bell Laboratories (Becker, 1994) and elsewhere would combine forces with (b) developments in data analysis (EDA, psychometrics, etc.) and (c) display and input technology (pen plotters, graphic terminals, digitizer tablets, the mouse, etc.). These developments would provide new paradigms, languages and software packages for expressing statistical ideas and implementing data graphics. In turn, they would lead to an explosive growth in new visualization methods and techniques.

Other themes began to emerge, mostly as initial suggestions: (a) various novel visual representations of multivariate data (Andrews’ (1972) Fourier function plots, Chernoff (1973) faces, star plots, clustering and tree representations); (b) the development of various dimension-reduction techniques (biplot (Gabriel, 1971), multidimensional scaling, correspondence analysis), providing visualization of multidimensional data in a 2D approximation; (c) animations

of a statistical process; and (d) perceptually-based theory and experiments related to how graphic attributes and relations might be rendered to better convey the data visually.

By the close of this period, the first exemplars of modern GIS and interactive systems for 2D and 3D statistical graphics would appear. These would set goals for future development and extension.

## 2.8 1975–present: High-D, interactive and dynamic data visualization

During the last quarter of the 20<sup>th</sup> century data visualization has blossomed into a mature, vibrant and multi-disciplinary research area, as may be seen in this Handbook, and software tools for a wide range of visualization methods and data types are available for every desktop computer. Yet, it is hard to provide a succinct overview of the most recent developments in data visualization, because they are so varied, have occurred at an accelerated pace, and across a wider range of disciplines. It is also more difficult to highlight the most significant developments, that may be seen as such in a subsequent history focusing on this recent period.

With this disclaimer, a few major themes stand out:

- the development of highly interactive statistical computing systems. Initially, this meant largely command-driven, directly programmable systems (APL, S), as opposed to compiled, batch processing;
- new paradigms of direct manipulation for visual data analysis (linking, brushing ([Becker and Cleveland, 1987](#)), selection, focusing, etc.);
- new methods for visualizing high-dimensional data (the grand tour ([Asimov, 1985](#)), scatterplot matrix ([Tukey and Tukey, 1981](#)), parallel coordinates plot ([Inselberg, 1985](#), [Wegman, 1990](#)), spreadplots ([Young, 1994a](#)), etc.);
- the invention (or re-invention) of graphical techniques for discrete and categorical data;
- the application of visualization methods to an ever-expanding array of substantive problems and data structures, and
- substantially increased attention to the cognitive and perceptual aspects of data display.

These developments in visualization methods and techniques arguably depended on advances in theoretical and technological infrastructure, perhaps more so than in previous periods. Some of these are:

- large-scale statistical and graphics software engineering, both commercial (e.g., SAS) and non-commercial (e.g., Lisp-Stat, the R project). These have often been significantly leveraged by open-source standards for information presentation and interaction (e.g., Java, Tcl/Tk);
- extensions of classical linear statistical modeling to ever wider domains (generalized linear models, mixed models, models for spatial/geographical data, and so forth).
- vastly increased computer processing speed and capacity, allowing computationally intensive methods (bootstrap methods, Bayesian MCMC analysis, etc.), access to massive data problems (measured in terabytes) and real-time streaming data. Advances in this area continue to press for new visualization methods.

From the early 1970s to mid 1980s, many of the advances in statistical graphics concerned static graphs for multidimensional quantitative data, designed to allow the analyst to see relations in progressively higher dimensions. Older ideas of dimension reduction techniques (prin-

principal component analysis, multidimensional scaling, discriminant analysis, etc.) led to generalizations of projecting a high-D dataset to “interesting” low-D views, as expressed by various numerical indices that could be optimized (projection pursuit) or explored interactively (grand tour).

The development of general methods for multidimensional contingency tables began in the early 1970s, with Leo Goodman (1970), Shelly Haberman (1973) and others (Bishop *et al.*, 1975) laying out the fundamentals of log-linear models. By the mid 1980s, some initial, specialized techniques for visualizing such data were developed (fourfold display (Fienberg, 1975), association plot (Cohen, 1980), mosaic plot (Hartigan and Kleiner, 1981) and sieve diagram (Riedwyl and Schüpbach, 1983)), based on the idea of displaying frequencies by area (Friendly, 1995). Of these, extensions of the mosaic plot (Friendly, 1994, 1999) have proved most generally useful, and are now widely implemented in a variety of statistical software, most completely in the `vcd` package (Meyer *et al.*, 2005) in R.

It may be argued that the greatest potential for recent growth in data visualization came from the development of dynamic graphic methods, allowing instantaneous and direct manipulation of graphical objects and related statistical properties. One early instance was a system for interacting with probability plots (Fowlkes, 1969) in realtime, choosing a shape parameter of a reference distribution and power transformations by adjusting a control. The first general system for manipulating high-dimensional data was PRIM-9, developed by Fishkeller, Friedman and Tukey (1974), and providing dynamic tools for Projecting, Rotating (in 3D), Isolating (identifying subsets) and Masking data in up to 9 dimensions. These were quite influential, but remained one-of-a-kind, “proof of concept” systems. By the mid 1980s, as workstations and display technology became cheaper and more powerful, desktop software for dynamic graphics became more widely available (e.g., MacSpin, Xgobi). Many of these developments to that point are detailed in the chapters of *Dynamic Graphics for Statistics* (Cleveland and McGill, 1988).

In the 1990s, a number of these ideas were brought together to provide more general systems for dynamic, interactive graphics, combined with data manipulation and analysis in coherent and extensible computing environments. The combination of all these factors was more powerful and influential than the sum of their parts. Lisp-Stat (Tierney, 1990) and its progeny (Arc, Cook and Weisberg (1999); ViSta, Young (1994b)), for example, provided an easily extensible object-oriented environment for statistical computing. In these systems, widgets (sliders, selection boxes, pick lists, etc.), graphs, tables, statistical models and the user all communicated through messages, acted upon by whomever was a designated “listener,” and had a method to respond. Most of the ideas and methods behind present day interactive graphics are described and illustrated in Young *et al.* (2006). Other chapters in this Handbook provide current perspectives on other aspects of interactive graphics.

### 3 Statistical historiography

As mentioned at the outset, this review is based on the information collected for the Milestones Project, which I regard (subject to some caveats) as a relatively comprehensive corpus of the significant developments in the history of data visualization. As such, it is of interest to consider what light modern methods of statistics and graphics can shed on this history, a self-referential question we call “statistical historiography” (Friendly, 2005). In return, this offers other ways



to view this history.

### 3.1 History as “data”

Historical events, by their nature, are typically discrete, but marked with dates or ranges of dates, and some description— numeric, textual, or classified by descriptors (who, what, where, amount, and so forth). Among the first to recognize that history could be treated as data and

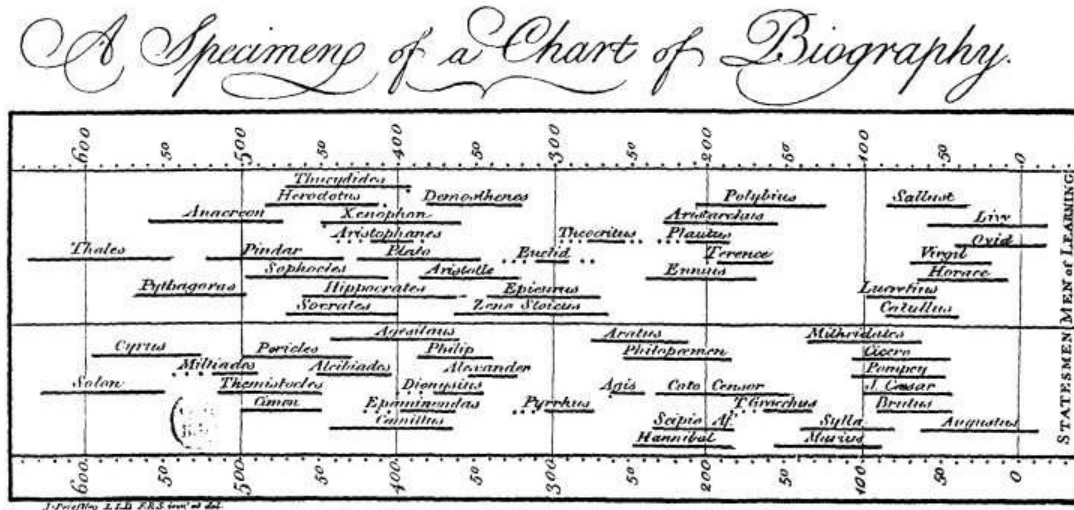


Figure 15: A specimen version of Priestley’s *Chart of Biography*. Source: Priestley (1765).

portrayed visually, Joseph Priestley (1765, 1769) developed the idea of depicting the lifespans of famous people by horizontal lines along a time scale. His enormous (2’ by 3’) and detailed *Chart of Biography* showed two thousand names from 1200 BC to 1750 AD by horizontal lines from birth to death, using dots at either end to indicate ranges of uncertainty. Along the vertical dimension, Priestley classified these individuals, e.g., as statesmen or men of learning. A small fragment of this chart is shown in Figure 15.

Priestley’s graphical representations of time and duration apparently influenced Playfair’s introduction of time-series charts and bar charts (Funkhouser, 1937, p. 280). But these inventions did not inspire the British statisticians of his day, as noted earlier; historical events and statistical facts were seen as separate, rather than as data arrayed along a time dimension. In 1885 at the Jubilee meeting of the Royal Statistical Society, Alfred Marshall (1885) argued that the causes of historical events could be understood by the use of statistics displayed by “historical curves” (time-series graphs): “I wish to argue that that the graphic method may be applied as to enable history to do this work better than it has hitherto” (p. 252). Maas and Morgan (2005) discuss these issues in more detail.

### 3.2 Analyzing Milestones data

The information collected in the Milestone Project is rendered in print and web forms as a chronological list, but is maintained as a relational database (historical items, references, images) in order to be able to work with it as “data.” The simplest analyses examine trends over

time. Figure 1 shows a density estimate for the distribution of 248 milestones items from 1500 to the present, keyed to the labels for the periods in history. The bumps, peaks and troughs all seem interpretable: note particularly the steady rise up to  $\sim 1880$ , followed by a decline through the “modern dark ages” to  $\sim 1945$ , then the steep rise up to the present. In fact, it is slightly surprising to see that the peak in the Golden Age is nearly as high as that at present, but this probably just reflects under-representation of the most recent events.<sup>23</sup>

Other historical patterns can be examined by classifying the items along various dimensions (place, form, content, and so forth). If we classify the items by place of development (Europe vs. North America, ignoring Other), interesting trends appear (Figure 16). The greatest peak in Europe around 1875–1880 coincided with a smaller peak in North America. The decline in Europe following the Golden Age was accompanied by an initial rise in North America, largely due to popularization (e.g., text books) and significant applications of graphical methods, then a steep decline as mathematical statistics held sway.

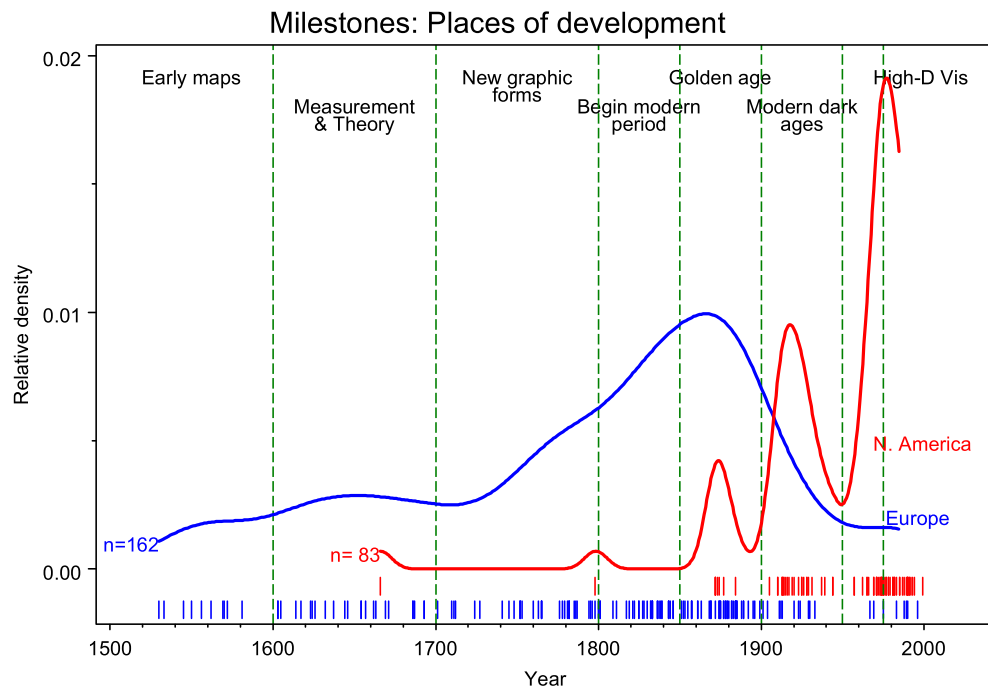


Figure 16: The distribution of milestone items over time, comparing trends in Europe and North America.

Finally, Figure 17 shows two mosaic plots for the milestones items classified by Epoch, Subject matter and Aspect. Subject was classed as having to do with human (e.g., mortality, disease), physical or mathematical characteristics of what was represented in the innovation. Aspect classed each item according to whether it was primarily map-based, a diagram or statis-

<sup>23</sup>Technical note: In this figure an optimal bandwidth for the kernel density estimate was selected (using the Sheather-Jones plugin estimate) for each series separately. The smaller range and sample size of the entries for Europe vs. North America gives a smaller bandwidth for the former, by a factor of about 3. Using a common bandwidth, fixed to that determined for the whole series (Figure 1) undersmooths the more extensive data on European developments and oversmooths the North American ones. The details differ, but most of the points made in the discussion about what was happening when and where hold.

tical innovation or a technological one. The left mosaic shows the shifts in Subject over time: Most of the early innovations concerned physical subjects, while the later periods shift heavily to mathematical ones. Human topics are not prevalent overall, but were dominant in the 19<sup>th</sup> century. The right mosaic, for Subject  $\times$  Aspect indicates that, unsurprisingly, map-based innovations were mainly about physical and human subjects, while diagrams and statistical ones were largely about mathematical subjects. Historical classifications clearly rely on more detailed definitions than described here, however, it seems reasonable to suggest that such analyses of history as “data” are a promising direction for future work.

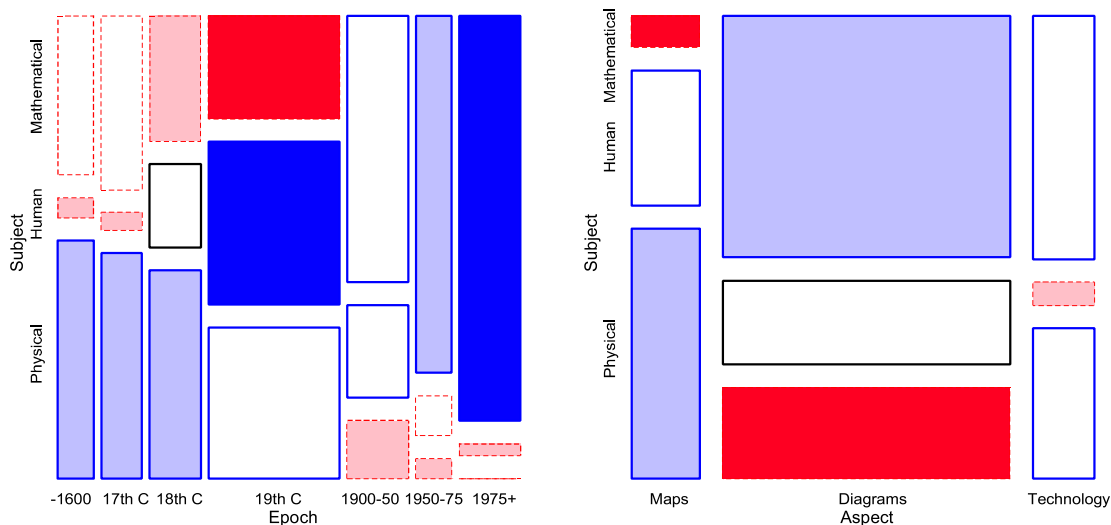


Figure 17: Mosaic plots for milestones items, classified by Subject, Aspect and Epoch.

### 3.3 What was he thinking?: Understanding through reproduction

Historical graphs were created using available data, methods, technology, and understanding current at the time. We can often come to a better understanding of intellectual, scientific, and graphical questions by attempting a re-analysis from a modern perspective.

Earlier, we showed Playfair’s time-series graph (Figure 7) of wages and prices, and noted that Playfair wished to show that workers were better off at the end of the period shown than at any earlier time. Presumably he wished to draw the reader’s eye to the narrowing of the gap between the bars for prices and the line graph for wages. Is this what you see?

What this graph shows directly is quite different than Playfair’s intention. It appears that wages remained relatively stable, while the price of wheat varied greatly. The inference that wages increased relative to prices is indirect and not visually compelling.

We cannot resist the temptation to give Playfair a helping hand here—by graphing the ratio of wages to prices (labor cost of wheat), as shown in Figure 18. But this would not have occurred to Playfair, because the idea of relating one time series to another by ratios (index numbers) would not occur for another half-century (due to Jevons). See [Friendly and Denis \(2005\)](#) for further discussion of Playfair’s thinking.

As another example, we give a brief account of an attempt to explore Galton’s discovery of regression and the elliptical contours of the bivariate normal surface, treated in more detail in



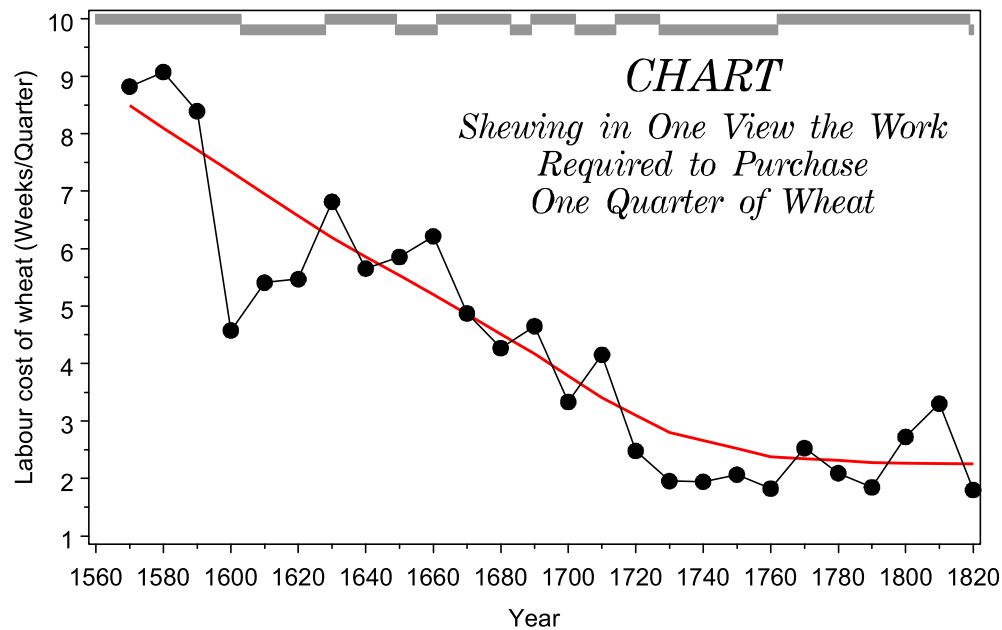


Figure 18: Redrawn version of Playfair’s time series graph showing the ratio of price of wheat to wages, together with a loess smoothed curve.

Friendly and Denis (2005). Galton’s famous graph showing these relations (Figure 19) portrays the joint frequency distribution of the height of children and the average height of their parents. It was produced from a “semi-graphic table” in which Galton averaged the frequencies in each set of four adjacent cells, drew iso-curves of equal smoothed value, and noted that these formed “concentric and similar ellipses.”

A literal transcription of Galton’s method, using contour curves of constant average frequency, and showing the curves of the means of  $y | x$  and  $x | y$  is shown in Figure 20. It is not immediately clear that the contours are concentric ellipses, nor that the curves of means are essentially linear and have horizontal and vertical tangents to the contours.

A modern data analyst following the spirit of Galton’s method might substitute a smoothed bivariate kernel density estimate for Galton’s simple average of adjacent cells. The result, using sunflower symbols to depict the cell frequencies, and a smoothed loess curve to show  $\mathcal{E}(y | x)$  is shown in Figure 21. The contours now *do* emphatically suggest concentric similar ellipses, and the regression line is near the points of vertical tangency. A reasonable conclusion from these figures is that Galton did not slavishly interpolate iso-frequency values as is done in the contour plot shown in Figure 20. Rather, he drew his contours to the smoothed data by eye and brain (as he had done earlier with maps of weather patterns), with knowledge that he could, as one might say today, trade some increase in bias for a possible decrease in variance, and so achieve a greater smoothing.

## 4 Final thoughts

This chapter is titled “A brief history ...” out of recognition that it is impossible to do full justice to the history of data visualization in such a short account. This is doubly so because I

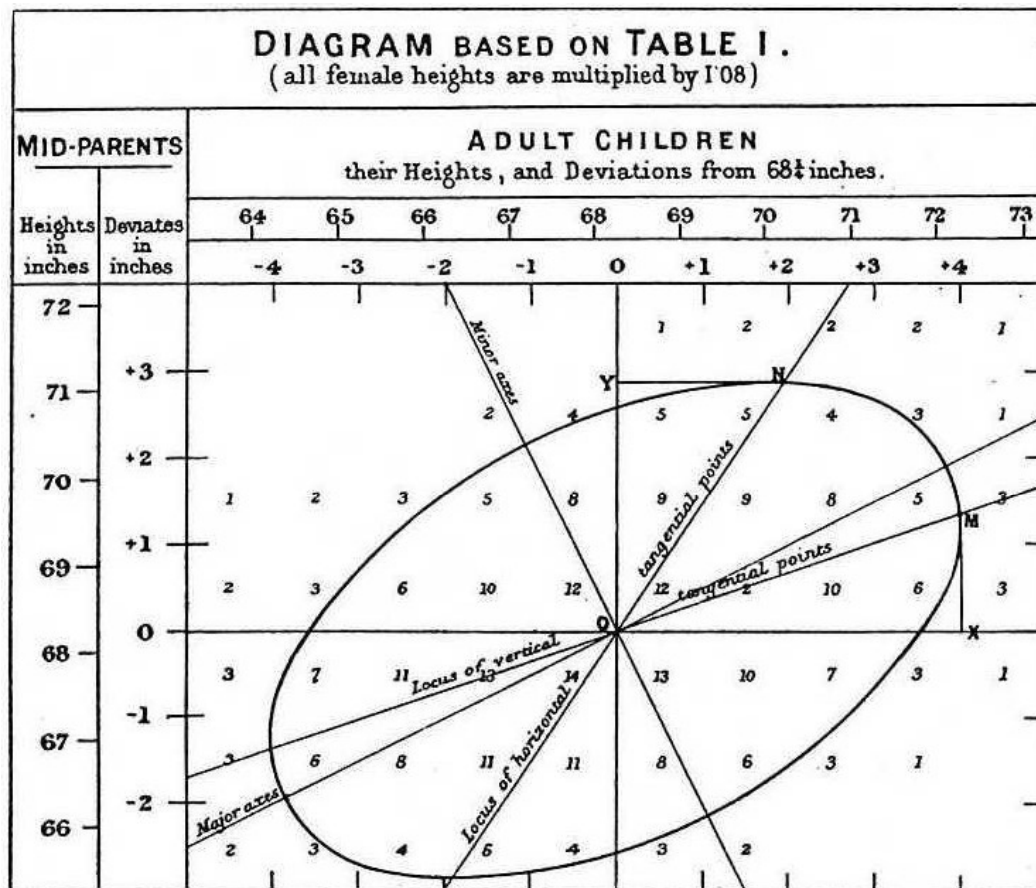


Figure 19: Galton's smoothed correlation diagram for the data on heights of parents and children, showing one ellipse of equal frequency. Source: (Galton, 1886, Plate X.).

have attempted to present a broad view spanning the many areas of application in which data visualization took root and developed. That being said, it is hoped that this overview will lead modern readers and developers of graphical methods to appreciate the rich history behind the latest hot new methods. As we have seen, almost all current methods have a much longer history than is commonly thought. Moreover, as I have surveyed this work and traveled to many libraries to view original works and read historical sources, I have been struck with the exquisite beauty and attention to graphic detail seen in many of these images, particularly those from the 19<sup>th</sup> century. We would be hard-pressed to recreate many of these today.

From this history one may also see that most of the innovations in data visualization arose from concrete, often practical goals: the need or desire to see phenomena and relationships in new or different ways. It is also clear that the development of graphic methods depended fundamentally on parallel advances in technology, data collection and statistical theory. Finally, I believe that the application of modern methods of data visualization to its own history, in this self-referential way I call “statistical historiography,” offers some interesting views of the past and challenges for the future.

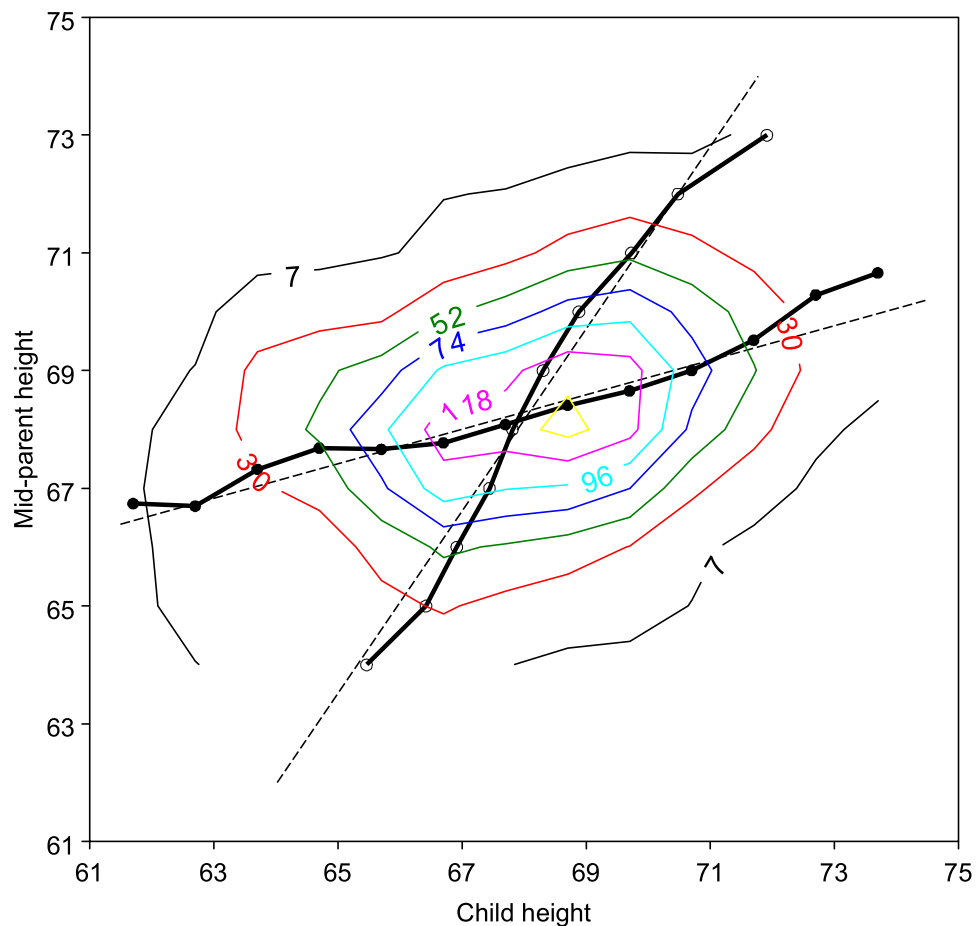


Figure 20: Contour plot of Galton's smoothed data, showing the curves of  $\bar{y} | x$  (filled circles, solid line),  $\bar{x} | y$  (open circles, solid line) and the corresponding regression lines (dashed).

## References

A few items in this reference list are identified by shelfmarks or call numbers in the following libraries: **BL**: British Library, London; **BNF**: Bibliothèque Nationale de France, Paris (Tolbiac); **ENPC**: École Nationale des Ponts et Chaussées, Paris; **LC**: Library of Congress; **SBB**: Staatsbibliothek zu Berlin.

Abbott, E. A. (1884). *Flatland: A Romance of Many Dimensions*. Cutchogue, NY: Buccaneer Books. (1976 reprint of the 1884 edition). 14

Andrews, D. F. (1972). Plots of high dimensional data. *Biometrics*, 28, 125–136. 23

Asimov, D. (1985). Grand tour. *SIAM Journal of Scientific and Statistical Computing*, 6(1), 128–143. 24

Ayres, L. P. (1919). *The War with Germany, A Statistical Summary*. Washington, D.C.: U.S. Government Printing Office. Commonly known as the *Ayres report*; reprinted: Arno Press, NY, 1979. 21

Baker, R. (1833). Report of the Leeds Board of Health. BL: 10347.ee.17(8). 12

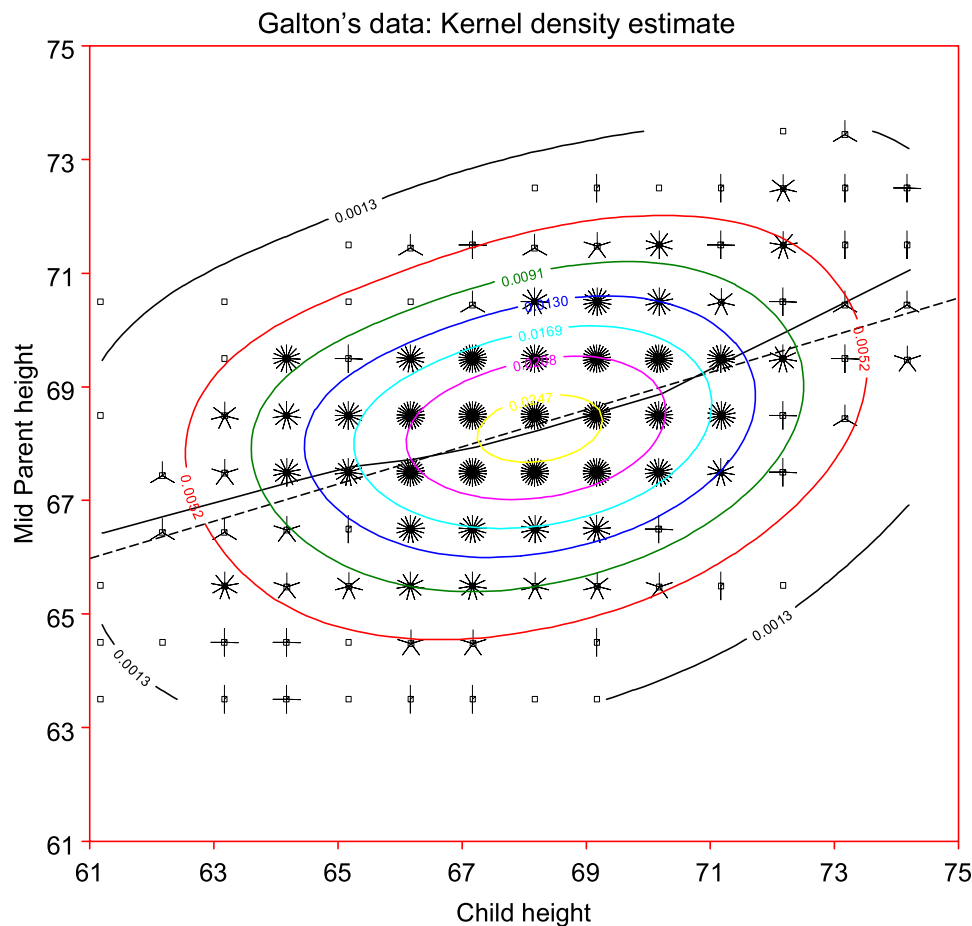


Figure 21: Bivariate kernel density estimate of Galton's data, using sunflower symbols for the data, and a smoothed loess curve for  $\mathcal{E}(y | x)$  (solid) and regression line (dashed).

Ball, W. W. R. (1908). *A Short Account of the History of Mathematics*. London: Macmillan & Co., 4th edn. (re-published in 1960, N.Y.: Dover). 6

Becker, R. A. (1994). A brief history of S. In P. Dirschedl and R. Ostermann, eds., *Computational Statistics*, (pp. 81–110). Heidelberg: Physica Verlag. 23

Becker, R. A. and Cleveland, W. S. (1987). Brushing scatterplots. *Technometrics*, 29, 127–142. 24

Berghaus, H. (1838). *Physikalischer Atlas*. Gotha: Justus Perthes. 2 vols., published 1845–48. 14

Bertin, J. (1967). *Sémiologie Graphique: Les diagrammes, les réseaux, les cartes*. Paris: Gauthier-Villars. 23

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press. 25

Bowley, A. L. (1901). *Elements of Statistics*. London: P. S. King and Son. 21, 22

- Buache, P. (1752). Essai de géographie physique. *Mémoires de L'Académie Royale des Sciences*, (pp. 399–416). BNF: Ge.FF-8816-8822. 7
- Chernoff, H. (1973). The use of faces to represent points in  $k$ -dimensional space graphically. *Journal of the American Statistical Association*, 68, 361–368. 23
- Cleveland, W. S. and McGill, M. E., eds. (1988). *Dynamic Graphics for Statistics*. Boca Raton, FL: CRC Press. 25
- Cohen, A. (1980). On the graphical display of the significant components in a two-way contingency table. *Communications in Statistics—Theory and Methods*, A9, 1025–1041. 25
- Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. New York: Wiley. 25
- Costelloe, M. F. P. (1915). Graphic methods and the presentation of this subject to first year college students. *Nebraska Blue Print*. 21
- Crome, A. F. W. (1785). *Über die Grösse and Bevölkerung der Sämtlichen Europäischen Staaten*. Leipzig: Weygand. 8
- Dahmann, D. C. (2001). Presenting the nation's cultural geography [in census atlases]. At American Memory: Historical Collections for the National Digital Library Internet. [Http://memory.loc.gov/ammem/gmdhtml/census.html](http://memory.loc.gov/ammem/gmdhtml/census.html). 20
- de Fourcroy, C. (1782). *Essai d'une table poléométrique, ou amusement d'un amateur de plans sur la grandeur de quelques villes*. Paris: Dupain-Triel. 8
- de Witt, J. (1671). *The Worth of Life Annuities Compared to Redemption Bonds*. Leiden: n.p. 6
- du Carla-Boniface, M. (1782). Expression des nivellements; ou, méthode nouvelle pour marquer sur les cartes terrestres et marines les hauteurs et les configurations du terrain. In François de Dainville, “From the Depths to the Heights,” translated by Arthur H. Robinson, *Surveying and Mapping*, 1970, 30:389–403, on page 396. 7
- Dupin, C. (1826). *Carte figurativ de l'instruction populaire de la France*. Jobard. BNF: Ge C 6588 (Funkhouser (1937, p. 300) incorrectly dates this as 1819). 10
- Eells, W. C. (1926). The relative merits of circles and bars for representing component parts. *Journal of the American Statistical Association*, 21, 119–132. 22
- Farebrother, R. W. (1999). *Fitting Linear Relationships: A History of the Calculus of Observations 1750–1900*. New York: Springer. 2
- Ferguson, S. (1991). The 1753 carte chronographique of Jacques Barbeu-Dubourg. *Princeton University Library Chronicle*, 52, 190–230. 8
- Fienberg, S. E. (1975). Perspective Canada as a social report. Tech. rep., Department of Applied Statistics, University of Minnesota. Unpublished paper. 25
- Fishkeller, M. A., Friedman, J. H., and Tukey, J. W. (1974). PRIM-9: An interactive multidimensional data display and analysis system. Tech. Rep. SLAC-PUB-1408, Stanford Linear Accelerator Center, Stanford, CA. 23, 25
- Fowlkes, E. B. (1969). User's manual for a system for interactive probability plotting on Graphic-2. Tech. rep., Bell Laboratories. 23, 25

- Friendly, M. (1994). Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89, 190–200. 12, 18, 25
- Friendly, M. (1995). Conceptual and visual models for categorical data. *The American Statistician*, 49, 153–160. 25
- Friendly, M. (1999). Extending mosaic displays: Marginal, conditional, and partial views of categorical data. *Journal of Computational and Graphical Statistics*, 8(3), 373–395. 25
- Friendly, M. (2000). Re-Visions of Minard. *Statistical Computing & Statistical Graphics Newsletter*, 11(1), 1, 13–19. 12
- Friendly, M. (2002). Visions and re-visions of Charles Joseph Minard. *Journal of Educational and Behavioral Statistics*, 27(1), 31–52. 14
- Friendly, M. (2005). Milestones in the history of data visualization: A case study in statistical historiography. In C. Weihs and W. Gaul, eds., *Classification: The Ubiquitous Challenge*, (pp. 34–52). New York: Springer. 25
- Friendly, M. and Denis, D. (2000). The roots and branches of statistical graphics. *Journal de la Société Française de Statistique*, 141(4), 51–60. (published in 2001). 20
- Friendly, M. and Denis, D. (2005). The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*, 41(2), 103–130. 11, 22, 28, 29
- Friendly, M. and Kwan, E. (2003). Effect ordering for data displays. *Computational Statistics and Data Analysis*, 43(4), 509–539. 6, 20
- Friis, H. R. (1974). Statistical cartography in the United States prior to 1870 and the role of Joseph C. G. Kennedy and the U.S. Census Office. *American Cartographer*, 1, 131–157. 2
- Frisius, R. G. (1533). *Libellus de locorum describendorum ratione*. Antwerp. 4
- Funkhouser, H. G. (1936). A note on a tenth century graph. *Osiris*, 1, 260–262. 2, 3, 4
- Funkhouser, H. G. (1937). Historical development of the graphical representation of statistical data. *Osiris*, 3(1), 269–405. Reprinted Brugge, Belgium: St. Catherine Press, 1937. 2, 12, 14, 18, 26
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal components analysis. *Biometrics*, 58(3), 453–467. 23
- Galton, F. (1863). *Meteorographica, or, Methods of Mapping the Weather*. London: Macmillan. BL: Maps.53.b.32. 16
- Galton, F. (1881). On the construction of isochronic passage charts. *Proceedings of the Geographical Section of the British Association*, n.v.(XI), 657, 704. Read Sept. 1, 1881; also published: *Roy. Geog. Soc. Proc.*, 1881, 657–658. 16
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15, 246–263. 16, 30
- Galton, F. (1894). Results derived from the natality table of Körösi by employing the method of contours or isogens. *Journal of the Royal Statistical Society*, 57(4), 702–708. 16
- Giffen, R. (1899). The excess of imports. *Journal of the Royal Statistical Society*, 62(1), 1–82. 21



- Gilbert, E. W. (1958). Pioneer maps of health and disease in England. *Geographical Journal*, 124, 172–183. 11, 12
- Goodman, L. A. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Association*, 65, 226–256. 25
- Graunt, J. (1662). *Natural and Political Observations Mentioned in a Following Index and Made Upon the Bills of Mortality*. London: Martin, Allestry, and Dicas. 6
- Guerry, A.-M. (1833). *Essai Sur La Statistique Morale de la France*. Paris: Crochard. English translation: Hugh P. Whitt and Victor W. Reinking, Lewiston, N.Y. : Edwin Mellen Press, 2002. 10, 39
- Guerry, A.-M. (1864). *Statistique morale de l'Angleterre comparée avec la statistique morale de la France, d'après les comptes de l'administration de la justice criminelle en Angleterre et en France, etc.* Paris: J.-B. Baillière et fils. BNF: GR FOL-N-319; SG D/4330; BL: Maps 32.e.34; SBB: Fe 8586; LC: 11005911. 11
- Haberman, S. J. (1973). The analysis of residuals in cross-classified tables. *Biometrics*, 29, 205–220. 25
- Hald, A. (1990). *A History of Probability and Statistics and their Application before 1750*. New York: John Wiley and Sons. 2, 5
- Halley, E. (1686). On the height of the mercury in the barometer at different elevations above the surface of the earth, and on the rising and falling of the mercury on the change of weather. *Philosophical Transactions*, (pp. 104–115). 6
- Halley, E. (1701). The description and uses of a new, and correct sea-chart of the whole world, shewing variations of the compass. London. 7, 14
- Hartigan, J. A. and Kleiner, B. (1981). Mosaics for contingency tables. In W. F. Eddy, ed., *Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface*, (pp. 268–273). New York, NY: Springer-Verlag. 25
- Haskell, A. C. (1919). *How to Make and Use Graphic Charts*. New York: Codex. 21
- Herschel, J. F. W. (1833). On the investigation of the orbits of revolving double stars. *Memoirs of the Royal Astronomical Society*, 5, 171–222. 9, 10
- Hertzprung, E. (1911). Publikationen des astrophysikalischen observatorium zu Potsdam. Num. 63. 21
- Hoff, H. E. and Geddes, L. A. (1959). Graphic recording before Carl Ludwig: An historical summary. *Archives Internationales d'Histoire des Sciences*, 12, 3–25. 2
- Hoff, H. E. and Geddes, L. A. (1962). The beginnings of graphic recording. *Isis*, 53, 287–324. Pt. 3. 2
- Howard, L. (1800). On a periodical variation of the barometer, apparently due to the influence of the sun and moon on the atmosphere. *Philosophical Magazine*, 7, 355–363. 8
- Inselberg, A. (1985). The plane with parallel coordinates. *The Visual Computer*, 1, 69–91. 24
- Jevons, W. S. (1863). A serious fall in the value of gold ascertained, and its social effects set fourth. London. 15

- Jevons, W. S. ([1879] 1958). Graphical method. In *Principles of Science: A Treatise on Logic and Scientific Method*, (pp. 492–496). New York: Dover, 3rd edn. First ed.: 1874; page numbers from 3rd Ed. Dover reprint (1958). 15
- Joint Committee on Standards for Graphic Presentation (1914). Preliminary report published for the purpose of inviting suggestions for the benefit of the committee. *Publications of the American Statistical Association*, 14(112), 790–797. 22
- Karsten, K. G. (1925). *Charts and Graphs. An Introduction to Graphic Methods in the Control and Analysis of Statistics*. New York: Prentice Hall. 21
- Klein, J. L. (1997). *Statistical Visions in Time: A History of Time Series Analysis, 1662-1938*. Cambridge, UK: Cambridge University Press. 2, 6, 21
- Kruskal, W. (1977). Visions of maps and graphs. In *Proceedings of the International Symposium on Computer-Assisted Cartography, Auto-Carto II*, (pp. 27–36). 1975. 2
- Lallemand, C. (1885). *Les abaques hexagonaux: Nouvelle méthode générale de calcul graphique, avec de nombreux exemples d'application*. Paris: Ministère des travaux publics, Comité du nivellement général de la France. 16
- Maas, H. and Morgan, M. S. (2005). Timing history: The introduction of graphical analysis in 19<sup>th</sup> century british economics. *Revue d'Histoire des Sciences Humaines*, 7, 97–127. 13, 26
- Marey, E.-J. (1878). *La Méthode Graphique dans les Sciences Expérimentales et Principalement en Physiologie et en Médecine*. Paris: G. Masson. 21
- Marshall, A. (1885). On the graphic method of statistics. *Journal of the Royal Statistical Society*, Jubilee volume, 251–260. 26
- Maunder, E. W. (1904). Note on the distribution of sun-spots in heliographic latitude, 1874 to 1902. *Royal Astronomical Society Monthly Notices*, 64, 747–761. 21
- Meyer, D., Zeileis, A., and Hornik, K. (2005). *vcd: Visualizing Categorical Data*. R package version 0.9-5. 25
- Minard, C. J. (1826). *Projet de canal et de chemin de fer pour le transport de paves à Paris, précédé d'un tableau des progrès de la dépense du pavé de Paris pendant les deux derniers siècles*. Paris: A. Pihan Delaforest. ENPC: 4.5065/C289. 14
- Minard, C. J. (1844). Tableaux figuratifs de la circulation de quelques chemins de fer. lith. (n.s.). ENPC: 5860/C351, 5299/C307. 12
- Minard, C. J. (1861). *Des Tableaux Graphiques et des Cartes Figuratives*. Paris: E. Thunot et Cie. ENPC: 3386/C161; BNF: V-16168. 14
- Moseley, H. (1913). The high frequency spectra of the elements. *Philosophical Magazine*, (pp. 1024–). (continued, 1914, pp. 703–). 22
- Mouat, F. J. (1885). History of the Statistical Society of London. *Journal of the Royal Statistical Society*, Jubilee volume, 14–62. 13
- Nautonier, G. L. (1602–1604). *Mecometrie de l'eymant, c'est a dire la maniere de mesurer les longitudes par le moyen de l'eymant*. Paris: n.p. BL: 533.k.9. 14
- Nightingale, F. (1857). *Mortality of the British Army*. London: Harrison and Sons. 15

- Oresme, N. (1482). *Tractatus de latitudinibus formarum*. Padova. BL: IA 3Q024. 4
- Oresme, N. (1968). *Nicole Oresme and the Medieval Geometry of Qualities and Motions: A Treatise on the Uniformity and Difformity Known as Tractatus de Configurationibus Qualitatum et Motuum*. Madison WI: University of Wisconsin Press. Tr.: M. Clagget. 4
- Palsky, G. (1996). *Des Chiffres et des Cartes: Naissance et développement de la Cartographie Quantitative Française au XIX<sup>e</sup> siècle*. Paris: Comité des Travaux Historiques et Scientifiques (CTHS). 2, 7, 10, 14, 15, 18
- Pearson, E. S., ed. (1978). *The History of Statistics in the 17th and 18th Centuries Against the Changing Background of Intellectual, Scientific and Religious Thought*. London: Griffin & Co. Ltd. Lectures by Karl Pearson given at University College London during the academic sessions 1921–1933. 2
- Pearson, K. (1914–1930). *The life, letters and labours of Francis Galton*. Cambridge: University Press. 16, 17
- Pearson, K. (1920). Notes on the history of correlation. *Biometrika*, 13(1), 25–45. 16
- Peddle, J. B. (1910). *The Construction of Graphical Charts*. New York: McGraw-Hill. 21
- Perozzo, L. (1880). Della rappresentazione grafica di una collettività di individui nella successione del tempo. *Annali di Statistica*, 12, 1–16. BL: S.22. 14
- Petty, W. (1665). *The economic writings of Sir William Petty: together with the observations upon the bills of mortality*. Cambridge: The University Press. C. H. Hall (ed.) (more probably by Captain John Graunt). 6
- Playfair, W. (1786). *Commercial and Political Atlas: Representing, by Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure, and Debts of England, during the Whole of the Eighteenth Century*. London: Corry. Re-published in Wainer, H. and Spence, I. (eds.), *The Commercial and Political Atlas and Statistical Breviary*, 2005, Cambridge University Press, ISBN 0-521-85554-3. 8
- Playfair, W. (1801). *Statistical Breviary; Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe*. London: Wallis. Re-published in Wainer, H. and Spence, I. (eds.), *The Commercial and Political Atlas and Statistical Breviary*, 2005, Cambridge University Press, ISBN 0-521-85554-3. 8, 14
- Playfair, W. (1821). Letter on our agricultural distresses, their causes and remedies; accompanied with tables and copperplate charts shewing and comparing the prices of wheat, bread and labour, from 1565 to 1821. BL: 8275.c.64. 8, 10
- Porter, T. M. (1986). *The Rise of Statistical Thinking 1820–1900*. Princeton, NJ: Princeton University Press. 2
- Priestley, J. (1765). A chart of biography. London. BL: 611.l.19. 8, 26
- Priestley, J. (1769). *A New Chart of History*. London: Thomas Jeffreys. BL: Cup.1250.e.18 (1753). 8, 26
- Quetelet, A. (1831). *Recherches sur le Penchant au Crime aux Différens Âges*. Brussels: Hayez. English translation by Sawyer F. Sylvester. Cincinnati, OH: Anderson Publishing Company, 1984. 10

- Quetelet, A. (1835). *Sur l'homme et le développement de ses facultés, ou Essai d'une physique sociale*. Paris: Bachelier. English translation, by R. Knox and T. Smilbert, *A Treatise on Man and the Development of his Faculties*. Edinburgh, 1842; New York: Burt Franklin, 1968. 10
- Riddell, R. C. (1980). Parameter disposition in pre-Newtonian planetary theories. *Archives Hist. Exact Sci.*, 23, 87–157. 2
- Riedwyl, H. and Schüpbach, M. (1983). Siebdiagramme: Graphische darstellung von kontingenztafeln. Tech. Rep. 12, Institute for Mathematical Statistics, University of Bern, Bern, Switzerland. 25
- Robinson, A. H. (1982). *Early Thematic Mapping in the History of Cartography*. Chicago: University of Chicago Press. 2, 8
- Royston, E. (1970). Studies in the history of probability and statistics, III. a note on the history of the graphical presentation of data. *Biometrika*, (pp. 241–247). 43, Pts. 3 and 4 (December 1956); reprinted In *Studies in the History Of Statistics and Probability Theory*, eds. E. S. Pearson and M. G. Kendall, London: Griffin. 2
- Scheiner, C. (1626–1630). *Rosa Ursina sive Sol ex Admirando Facularum & Macularum Suarum Phoenomeno Varius*. Bracciano, Italy: Andream Phaeum. BL: 532.l.6. 5
- Smith, W. (1815). *A delineation of the strata of England and Wales, with part of Scotland; exhibiting the collieries and mines, the marshes and fenlands originally overflowed by the sea, and the varieties of soil according to the substrata, illustrated by the most descriptive names*. London: John Cary. BL: Maps 1180.(19). 9
- Snow, J. (1855). *On the Mode of Communication of Cholera*. London: (n.p.), 2nd edn. 12
- Sobel, D. (1996). *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York: Penguin. 5
- Spence, I. and Garrison, R. F. (1993). A remarkable scatterplot. *The American Statistician*, 47(1), 12–19. 21
- Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Harvard University Press. 2
- Tartaglia, N. F. (1556). *General Trattato di Numeri et Misura*. Venice: Vinegia. BL: 531.n.7-9; 47.e.4. 4
- Tierney, L. (1990). *LISP-STAT: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics*. New York: John Wiley and Sons. 25
- Tilling, L. (1975). Early experimental graphs. *British Journal for the History of Science*, 8, 193–213. 2
- Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press. 3, 5, 10, 14
- Tufte, E. R. (1997). *Visual Explanations*. Cheshire, CT: Graphics Press. 5, 6
- Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, 33, 1–67 and 81. 23

- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley. Chapters from what would become this book were circulated privately 1971–75. 23
- Tukey, P. A. and Tukey, J. W. (1981). Graphical display of data sets in 3 or more dimensions. In V. Barnett, ed., *Interpreting Multivariate Data*. Chichester, U.K.: Wiley and Sons. 24
- Vauthier, L. L. (1874). Note sur une carte statistique figurant la répartition de la population de Paris. *Comptes Rendus des Séances de L'Académie des Sciences*, 78, 264–267. ENPC: 11176 C612. 14
- von Huhn, R. (1927). A discussion of the Eells' experiment. *Journal of the American Statistical Association*, 22, 31–36. 22
- von Humboldt, A. (1817). Sur les lignes isothermes. *Annales de Chimie et de Physique*, 5, 102–112. 14
- Wainer, H. (2005). *Graphic Discovery: A Trout in the Milk and Other Visual Adventures*. Princeton, NJ: Princeton University Press. 16
- Wainer, H. and Velleman, P. F. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology*, 52, 305–335. 2
- Wallis, H. M. and Robinson, A. H. (1987). *Cartographical Innovations: An International Handbook of Mapping Terms to 1900*. Tring, Herts: Map Collector Publications. 2
- Warne, F. J. (1916). *Warne's Book of Charts, A Special Feature of Warne's Elementary Course in Chartography*. Washington, D.C.: F. J. Warne. 3 p. l., 106 charts. 31 x 41 cm. 21
- Washburne, J. N. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Psychology*, 18, 361–376, 465–476. 22
- Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411), 664–675. 24
- Westergaard, H. (1932). *Contributions to the History of Statistics*. London: P. S. King & Son. 17
- Whitt, H. (2002). Inventing sociology: André-Michel Guerry and the *Essai Sur La Statistique Morale de la France*. In [Guerry \(1833\)](#), (pp. ix–xxxvii). English translation: Hugh P. Whitt and Victor W. Reinking, Lewiston, N.Y. : Edwin Mellen Press, 2002. 11
- Young, F. W. (1994a). ViSta: The visual statistics system. Tech. Rep. RM 94-1, L.L. Thurstone Psychometric Laboratory, UNC. 24
- Young, F. W. (1994b). ViSta: The visual statistics system. Tech. Rep. RM 94-1, L.L. Thurstone Psychometric Laboratory, UNC. 25
- Young, F. W., Valero-Mora, P., and Friendly, M. (2006). *Visual Statistics: Seeing Data with Dynamic Interactive Graphics*. New York: Wiley. 25
- Zeuner, G. (1869). *Abhandlungen aus der mathematischen statistik*. Leipzig. BL: 8529.f.12. 14

## Index

- Abbott, Edwin, 14
- age pyramid, 18
- analysis
  - discriminant, 25
  - principal component, 25
- arithmetic
  - political, 5, 6
- atlas
  - cartographic, 4
  - statistical, 17, 20
- atomic number, 21
- Baker, Robert, 12
- bar chart, 8, 20, 26
  - divided, 12
- Barbeu-Dubourg, Jacques, 7
- Bell Laboratories, 23
- Benzécri, Jean-Paul, 23
- Bertin, Jacques, 23
- bilateral frequency polygon, 18
- bills of mortality, 6
- biplot, 23
- bivariate normal surface, 28
- bootstrap, 24
- Bowley, Arthur, 21, 22
- boxplot, 23
- Brahe, Tycho, 4–6, 16
- brushing, 24
- Buache, Phillippe, 7
- C, 23
- camera obscura, 4
- chart
  - bar, 8, 12, 20, 26
  - control, 21
  - Gantt, 21
  - isochronic, 16
  - pie, 8
  - time-series, 26
- Cheysson, Émile, 15, 17
- clustering, 23
- continuous distribution function, 6
- coordinate
  - parallel, 16
- coordinate paper, 8
- coordinate systems, 6
- correlation, 16
- coxcomb, 14
- Crome, August F.W., 8
- curve
  - contour, 29
  - cumulative frequency, 21
  - freehand, 21
  - frequency, 21
  - historical, 26
  - isodic, 16
  - isogen, 16
  - loess, 29
- curve fitting, 8
- Cusa, Nicolas of, 4
- data
  - multivariate, 23
- de Fourcroy, Charles, 8
- de Witt, Jan, 6
- density estimation, 29
- Descartes, René, 5
- diagram
  - alignment, 16
  - butterfly, 21
  - cartogram, 8, 18
  - Chernoff faces, 23
  - circle, 18
  - contour, 14, 16
  - coxcomb, 14
  - divided circle, 14
  - geometric, 3
  - Hertzprung-Russell, 21
  - historical, 21
  - nomogram, 16
  - pie, 18
  - rose, 14, 18
  - sieve, 25
  - statistical, 18
  - stereogram, 14
  - tree, 23
- dimension



- 3, 16
- display
  - fourfold, 25
  - trellis, 16
  - two-way table, 23
- du Carla-Boniface, Marcellin, 7
- Dupin, Charles, 10
- dynamic graphics, 25
- Fermat, Pierre, 5
- figure
  - geometric, 8
- flow map, 14, 18
- FORTRAN, 23
- Friendly, Michael, 25
- Galilei, Galileo, 5
- Galton, Francis, 16, 17, 28–29
- Gannett, Henry, 18
- Gauss, Carl Friedrich, 14
- Gemma-Frisius, Reginer, 4
- Giffen, Robert, 21
- Golden Age, 13, 27
- Goodman, Leo, 25
- grand tour, 24
- graph
  - age pyramid, 18
  - bilateral polygon, 20
  - circle, 8
  - hanging rootogram, 23
  - high-resolution, 23
  - line, 5, 8
  - time-series, 3, 8, 26, 28
- graph paper, 8
- Graunt, John, 5, 6
- grid
  - coordinate, 7
  - hexagonal, 16
- Guerry, André-Michel, 10, 14
- Haberman, Shelly, 25
- Halley, Edmund, 7
- Harrison, John, 5
- Herschel, William, 9
- Huygens, Christiaan, 6
- interactive graphics, 25
- interpolation, 8
- isoline, 16
- Java, 24
- Jevons, William Stanley, 13
- Lallemand, Charles, 16
- Lambert, Johann, 8
- Laplace, Pierre Simon, 14
- le Blon, Jacob, 8
- Levasseur, Pierre Émile, 17
- life table, 6
- line
  - contour, 7
  - isogon, 7
  - isoline, 7, 16
  - slope as rate, 8
  - timeline, 7
- line graph
  - 1D, 5
- linking, 24
- Lisp-Stat, 24, 25
- lithography, 8
- longitude, 6
- MacSpin, 25
- map
  - anamorphic, 16
  - cartography, 7
  - choropleth, 10
  - contour, 7
  - disease, 12
  - epidemiological, 15
  - flow, 14, 18
  - geological, 9
  - shaded, 20
  - thematic, 7, 14
  - topographic, 7
  - weather, 6, 17
- Marey, Étienne Jules, 21
- Marshall, Alfred, 26
- matrix
  - scatterplot, 24
- Maunder, E. W., 21
- measurement
  - physical, 5
- measurement error, 7

- Minard, Charles Joseph, 12, 14
- model
  - generalized linear, 24
  - mixed, 24
- mosaic plot, 12, 25, 28
- Moseley, Henry, 22
- multidimensional scaling, 23, 25
- Nightingale, Florence, 14
- nomogram, 16
- Oresme, Nicole, 4
- Ortelius, Abraham, 4
- parallel coordinates plot, 24
- Pascal, Blaise, 5
- Pearson, Karl, 16
- Perozzo, Luigi, 14
- Petermann, Augustus, 12
- Petty, William, 5, 6
- pie chart, 8
- Playfair, William, 8, 26, 28
- plot
  - association, 25
  - bivariate, 6
  - Fourier function, 23
  - log-log, 16, 21
  - mosaic, 12, 25, 28
  - parallel coordinate, 18
  - probability, 25
  - star, 18, 23
  - stem-leaf, 23
- political arithmetic, 6
- Priestley, Joseph, 8, 26
- PRIM-9, 25
- printing
  - three-colour, 8
- probability theory, 6
- projection
  - axonomic, 14
- Ptolemy, Claudius, 3, 6
- Quetelet, Adolphe, 11, 14, 17
- R, 24, 25
- regression, 16, 29
- Rheticus, Georg, 4
- Royal Statistical Society, 13, 26
- SAS, 24
- scaling
  - multidimensional, 23, 25
- scatterplot matrix, 24
- Scheiner, Christopher, 5
- Schwabe, Hermann, 17
- Senefelder, Aloys, 8
- sieve diagram, 25
- small multiple, 18
- small multiples, 5
- Smith, William, 9
- smoothing, 21, 29
- Snow, John, 12, 15
- spreadplot, 24
- star plot, 18, 23
- statistical atlas, 17
- statistics
  - demographic, 5
  - moral, 10
- stem-leaf plot, 23
- stereogram, 14
- stratigraphic geology, 9
- symbol
  - sunflower, 29
- table
  - graphical, 8
  - life, 6
  - semi-graphic, 29
- tableau-graphique, 12
- Tcl/Tk, 24
- thematic cartography, 7
- time-series, 8
  - multiple, 14
- timeline, 7, 26
- tour
  - grand, 24
- triangulation, 4
- Tufte, Edward, 5
- Tukey, John W., 23
- two-way table, 23
- UNIX, 23
- van Langren, Michael F., 5

ViSta, 25  
visual thinking, 6  
von Mayr, Georg, 17  
  
Walker, Francis A., 18  
wall quadrant, 4  
weather map, 6, 17  
  
Xgobi, 25  
  
Zeuner, Gustav, 14

# Nightingale

Journal of the  
Data Visualization Society

[Home](#)
[Topics in DataViz](#)
[How To](#)
[Community](#)
[About](#)
[SUBSCRIBE](#)

## TOPICS IN DATAVIZ

## REVIEW: Data + Diversity-Exploring the Data Visualizations of W.E.B. Du Bois

Pearle Lundeen • July 22, 2021



Some of W. E. B. DuBois's original charts in the background and their #DuBoisChallenge recreations.

I identify as “new to data.” So, when I attended [Tableau’s February Data + Diversity talk](#), I was totally sending out imposter syndrome vibes. In the Zoom chat, I made a bar chart joke to throw the others off my scent, while the rest worked on the obligatory metre-long chain of “Hello from [place name here]!” messages.

The talk was titled Data + Diversity: Exploring the Data Visualizations of W.E.B. Du Bois. Facilitated by writer and adjunct professor at Columbia University, Allen Hillery, the webinar explored the charts of W. E. B. Du Bois two ways: Jason Forrest brought us through a detailed history of Du Bois’ visualizations, from concept through development, and their presentation at the 1900s World’s Fair. In the second half, Anthony J. Starks took us through his efforts to digitally recreate some of the originals. This is my review of that webinar.

W. E. B. Du Bois was an African American sociologist, historian, and civil rights activist who lived during the turn of the 20th century. Jason’s presentation focused on his 1900 Paris World’s Fair Exhibition, titled [The Exhibit of American Negroes](#).

I intended to leave the talk on in the background while I fought with a belligerent waffle chart. But as Jason took us through Du Bois’ work, slide-by-slide, my attention to my own work fell away.

It was my first exposure to Du Bois’ work. What I saw floored me. I saw graphs and charts, but none like I’ve ever made—each was hand-drawn, hand-painted. Without computer assistance, he made precise lines, sharp edges, and powerful points. And they kept coming. In all, we looked at over thirty different visualizations, each one

## From our sponsors



Its the gift that keeps on giving!

SUBSCRIBE  
NOW! \$40 for 2 issues

Just in time for the holidays!  
Give the gift of Nightingale Magazine!



[SIGN UP FOR THE 'GALE](#)

Get our monthly newsletter! It's a fun overview of our content and community

[Sign Up](#)



Photograph of the Exhibit of the American Negroes at the Paris Exposition, 1900. From the Library of Congress.

depicting a different aspect of life as a Black American in the early 20<sup>th</sup> century. By the end of the webinar I was in shock. The clarity in Du Bois' visualizations was devastating. Tears were rolling down my face.

I only had one thought in my head when it was over: *what just happened?*

For me, what happened was a layering of different contexts that, when combined, moved me like nothing before. First, was the context of the past: W. E. B. Du Bois' original Paris exhibition as it was presented in 1900. Jason's part of the presentation underlined how the impact of Du Bois' work was reliant on the context in which it was originally presented. He went through Du Bois' and his team's painstaking effort to hand-collect, hand-illustrate, and exhibit life as a Black American at

the turn of the 20<sup>th</sup> century. He described an ambitious struggle, a careful balance of time, money, prejudice, and people.

As Jason spoke, I must admit, I felt guilty. I was watching the presentation on a computer more powerful than what was used to send people to the moon – certainly more powerful than anything Du Bois had access to. It's a humbling thing, that moment when you realize you're taking something for granted.

The narrative Du Bois was trying to create was also ambitious. Jason explained that Du Bois was making a push to "...sway the world's elite by upending the stereotypes and presenting a modern, successful, and educated people." Equally disruptive was Du Bois' method of visualization. Jason explained that Du Bois was one of the first practitioners of interactive dataviz, as some of the charts invited visitors to touch and feel them. Jason's own journey through W. E. B. Du Bois's work brought him to the Library of Congress, where he was able to observe one of the original charts, allowing him to see the fingerprints and smudges left by exhibition attendees as they laid each on top of the other, to view information in layers.

Ambitious as it was, Du Bois' message was not without challenge. Though he received a gold award in Paris, the message fell flat back in the United States. It would take until 1964 for America to pass the Civil Rights Act, which legislated the prohibition of discrimination on the basis of race, color, religion, sex, or national origin. Du Bois' work fell into obscurity well before that, only being digitized in 2013. By the end of Jason's presentation, it was clear to me that Du Bois tried to effect positive change by bringing the narrative of a burgeoning Black America to a new audience, but his message was ultimately stymied by the limited reach of the Paris exhibition.

Where Jason's part of the presentation was an invitation to see the context of the past, Anthony Stark's part of the presentation served as an invitation to apply the technological contexts of today to Du Bois' work. Anthony recreated several of Du Bois' original visualizations, going into detail on how he used decksh (pronounced "deckshell"), a Domain Specific Language (DSL) package to create charts and other visualizations as presentation slides using the Go language. With decksh, Anthony brought some of Du Bois' hard copies into the digital era. He provided his datasets and code on github. He also invited us to share our recreations online. Anthony's presentation made for a low-barrier way to accept that invitation to interact with the data in a way that appeals to today's data practitioner.

Watching each presentation made me aware of the privilege of my own situation—I have access to resources Du Bois didn't have: a computer, ease of access to information, self-publishing, and self-promotion capabilities as simple and as broad-reaching as a tweet. Basically, I have access to things Du Bois only wished for. Jason read a reflection from *The Autobiography of W. E. B. Du Bois*, where Du Bois states how he regretted failing to advertise his efforts: *"I did the deed but failed to advertise it... in the long run Advertising without the Deed was the only lasting value."*

A lingering thought persisted in the back of my mind as I realized this: would the impact of Du Bois' work be any different if he had the same access to technology and social reach as any one of us in today's age?



## DVS Membership + PRINT!

Why not become a DVS General Member \*AND\* upgrade to receive a print subscription with your membership for \$25 (a \$15 discount). Patron Membership also includes print subscription to Nightingale Magazine.

[Learn more!](#)

It turned out that question was less hypothetical than I thought; Jason and Anthony made mention of the #DuBoisChallenge hashtag, and that people from all over the world were posting their recreations on Twitter. “What effect was Twitter having on the message of W. E. B. Du Bois,” was the question on my mind.

I found that the hashtag was introduced to Twitter in February at the same time as this presentation. This was no coincidence, as our facilitator, Allen Hillery, was also one of the people who started the hashtag. Curious to learn more about the origins of both, I asked Allen if he'd like to talk about the presentation and the #DuBoisChallenge. Here's what he said.

**Pearle:** How did the Tableau talk originate?

**Allen:** Sekou Tyler and I are on Tableau's Community Equity Task Force (CETF). This is a group that is committed to making sure that the Tableau Community is inclusive and recognizes opportunities for everyone to engage. For the Data and Diversity Talk, I proposed that we cover W.E.B. Du Bois as an intersection of Black History Month and data! W.E.B. Du Bois does not get enough credit for his contributions to the world of dataviz. It's my aim for that to change. This is also inline with the spirit of the Tableau CETF – to ensure everyone has a voice. We started organizing speakers for the event and the rest is history!

**Pearle:** How do you think the Tableau talk went?

**Allen:** It went well! There were a lot of comments from the audience during the talk about how inspired they were from the presentations. There were also a good number of people who did not realize that Du Bois had done such work. We are all impressed about the level of intricacy behind the visualizations themselves including the spiral graph and wraparound bar charts. We had about 2,000 people who registered and over 1,200 people in attendance!

**Pearle:** What were your takeaways from that experience?

**Allen:** I was very humbled and proud by the experience. I pitched the idea and believed it would be awesome for everyone to explore. Getting all the amazing feedback on how everyone learned a lot about Du Bois and the data visualizations is immensely fulfilling. It is important to highlight him and widen the narrative of data visualization from a diversity point of view as well as application. Du Bois's work had set the ground for sociology. He designed a framework to intrinsically measure and survey communities. The results of those surveys are represented in the visualizations.

**Pearle:** How did you and Sekou Tyler decide to put the #DuBoisChallenge on?

**Allen:** Sekou and I were chatting one weekend about doing something to honor W.E.B. Du Bois. His thought was for he and I to recreate Du Bois' data visualizations and share them on Twitter. Thinking about community data projects like #MakeoverMonday, #TidyTuesday, and #WorkOutWednesday, I proposed we start a community hashtag where everyone could join in the fun!

**Pearle:** What impact did the #DuBoisChallenge have on the dataviz community?

**Allen:** The #DuBoisChallenge brought a lot of learning and collaboration. Many were challenged and excited in recreating DuBois' visualizations. A lot of participants commented on how challenging some of them were, like the Du Bois spiral!





City and Rural Population. From Library of Congress.

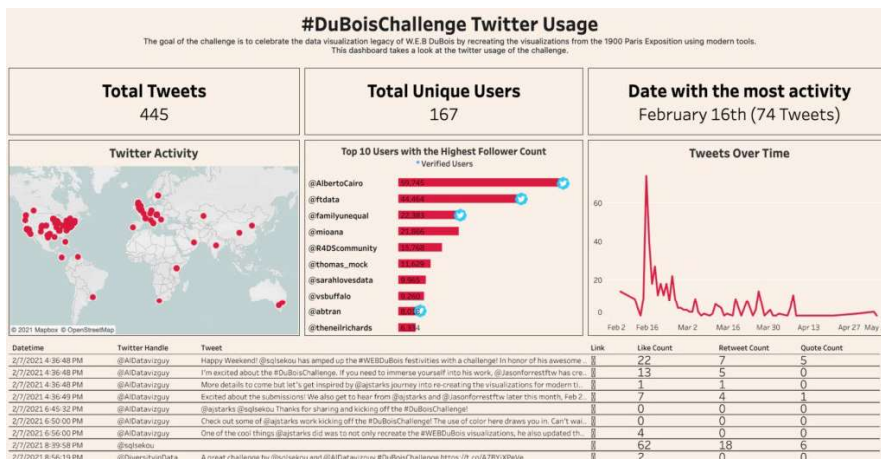
**Pearle:** What impact did the hashtag have on you?

**Allen:** I was able to work with some awesome people in the space to make this challenge happen. It's been great to work with Sekou and Anthony. I learned so much about Du Bois and the historical context of his work. Discovering that someone who looked like me made such a huge impact on the data visualization field has given me a sense of pride and confidence in my data journey. It's an honor and pleasure to get his work to be more of a household name in data cubicles.

As a data storyteller and someone who teaches the topic, I see Du Bois's work having the perfect ingredients for a well-balanced story. He uses data, emotion and credibility seamlessly to show the resilience of Black Americans in a post-slavery/Jim Crow era. In spite of all the challenges Black Americans endured socially and legally, they were showing an increase in the basic fabrics of life. This included education, home ownership, and income.

**Pearle:** Did anything unexpected come from the hashtag?

**Allen:** When we started the challenge the intent was to engage the data visualization community. As the challenge progressed, we gained interest from the humanities fields. Another great surprise was hearing from the R community via #TidyTuesday. So we ended up collaborating with another hashtag!! Sekou, Anthony, and I didn't expect the challenge to take off like it did. Sekou put together a [dashboard](#) to measure the engagement and reach of the challenge! We had 445 tweets around the challenge from February 2021 through May 2021. This sparked a follow up to the Du Bois challenge to see how the community would visualize the Twitter data!



#DuBoisChallenge 2021 Twitter Metrics dashboard. Credit: Sekou Tyler.

**Pearle:** How did Anthony Starks get involved in the #DuBoisChallenge?

**Allen:** I met Anthony at a Data Visualization Society event. We got to talking and he mentioned his Du Bois project! Anthony has not only recreated all of the Du Bois visualizations, he built his OWN tool to help him do so. As Sekou and I were building out the requirements for the challenge, I knew I had to contact Anthony. Thankfully he said yes. We owe Anthony so much for making this challenge what it is! He provided the data source as well as picked out the visualizations we should have in the challenge.



Some of W. E. B. Du Bois's original charts in the background and their #DuBoisChallenge recreations.

**Pearle:** What will you always remember when you think about the #DuBoisChallenge?

**Allen:** I will always remember the excitement and enthusiasm from the participants! I will remember the awesome people who either tweeted about the challenge or used their platforms to promote it. I will remember expanding our reach beyond the data visualization community and tapping into the humanities and sociology communities. We also have been able to talk at school events.

**Pearle:** What's next for the #DuBoisChallenge?

**Allen:** The challenge has been going in so many great directions! As I mentioned, we have been able to talk to classes and user groups about diversity and data, social justice, and leveraging social media to amplify your message. Another great path we're taking with the challenge is to build a course out of Du Bois' visualizations.

**Pearle:** What would you tell a person who wanted to do more along the lines of the #DuBoisChallenge?

**Allen:** I would tell them to participate in the challenge and understand the historical context. I would also encourage people to make Du Bois and his work part of the data visualization narrative because he is. One last thing I challenge folks to help me with is to find out more about the team that helped Du Bois build the visualizations! I would love to learn more about Du Bois's point person William Andrew Rogers!

Fast forward to today, and I'm still thinking about this webinar. Something enduring persists in my mind about

what happened on that February morning. When it was over, I exited the online webinar silently while the chat was a blur with many “THANK YOU!!!” and “Goodbye from [insert state here]” messages. I took off my headphones and sat back in my chair in silence to reflect on what I had gathered along the way. I had pieces of context. These pieces—the history of the past, Du Bois’ sense of opportunity lost, and the promise of today’s technology—they feel like the beginning of something, like a persistent nudge, early in the morning before the sun comes up. Some days it’s just too early and you roll over and go back to sleep. But some days you wake up, and that’s when you see small things change around you; the sky gets lighter, things get clearer, the birds start chirping.

Maybe that’s what this webinar is. It’s a gentle, but persistent, nudge to see. You don’t have to, of course. But if you do, maybe you’ll see things you might not have noticed before. And what you do with that knowledge can make a difference. At the end of the presentation, Jason Forrest left us with this thought: *“If you know the past, you are inspired to leverage it to create positive change.”* This webinar joins the idea of knowing the past with the tools of action in today’s digital world, and invites the audience to consider what positive change we can create when we use these tools to advertise—and amplify—the deed Du Bois started over 121 years ago.

#### Author profile



#### PEARLE LUNDEEN

Website

Pearle Lundeen is a Business Analyst in the Higher Education sector in Vancouver, British Columbia, Canada. When she’s not leading or supporting business process improvements and transformations, Pearle’s exploring her newfound passions: data, visualizations, and how they connect people.

#### Related entries

This author does not have any more posts.

#### CATEGORIES:

[COMMUNITY](#)
[DATA HUMANISM](#)
[HISTORIC DATAVIZ](#)
[REVIEWS](#)

TAGS: [data human](#), [data visualization](#), [dubois](#), [historic dataviz](#), [review](#), [webinar](#)

[← PREVIOUS ARTICLE](#)

**What Is the Bechdel Test and What Is Its Relevance to Today’s Film Industry?**

[NEXT ARTICLE →](#)

**You’ve Visualized Your Data, Now Actualize Your Promotion**

#### Related articles

MARY AVILES

#### REVIEW: A Guided Virtual Tour of ‘Data Visualization and the Modern Imagination’

Like inventors, explorers, and philosophers, data visualizers have honed their craft through experimentation, abject failure, breakthrough moments, pushing boundaries, and by standing on the shoulders..

JASON FORREST

#### Data Journalism in the study of W.E.B. Du Bois’s “The Negro Problem” (Part 2)

One of the most powerful examples of data visualization was made 118 years ago by an all-black team led by W.E.B. Du Bois only 37..

SIMRAN PARWANI

#### Early Career Corner: Meet Dzifa Amexo

I feel so lucky to be the Early Career Director at DVS: I get to talk to all sorts of people about data visualization and..

[Home](#)[Topics in Dataviz](#)[How To](#)[Community](#)[About](#)[SUBSCRIBE](#)

# Nightingale

Journal of the  
Data Visualization Society

[Home](#)[Topics in Dataviz](#)[How To](#)[Community](#)[About](#)[SUBSCRIBE](#)

## TOPICS IN DATAVIZ

### Beyond Nightingale: Being a Woman in Data Visualization

Stephanie Evergreen • July 15, 2019



Illustration by Jen Ray

If you conduct a quick internet search on “history of data visualization,” you’ll nearly always see Florence Nightingale included in the annals of history. Why? It’s not like a Nightingale Rose chart is easy to read, or a cinch to make, or even all that common.

One clue to the answer lies in the fact that she is most often the only woman on such lists.

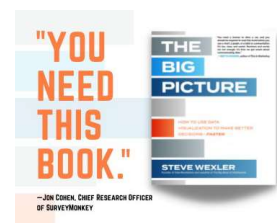
Many women know that when you are the only woman present on a panel full of men, two things are highly probable: (1) you are working in a male-dominated field, and (2) you are likely a token because *somebody* thought there had better be some diversity. **Florence Nightingale is our data visualization historical token female.**

History is written by cis white men. And history is upheld by them too, even among the data visualization crowd of today, who cull these lists of historically important figures and decide whose stories will be remembered, whose work will become “foundational.” At least initially, they, like me, perhaps took to Nightingale because they readily recognized her name from high school history books (written by — guess who) where her role in the war was recorded— though her Rose chart usually wasn’t printed.

But if her inclusion in the annals of viz history is really about how impactful or accomplished Nightingale was as a visualizer, we would see greater discussion of her **OTHER** charts and graphs, instead of reducing her contributions to a single image. We would learn that she was the first woman voted into the Royal Statistical Society. We would read about her as the equal to other “founding fathers” of the field. (RJ Andrews is expanding on this topic, which [you can find here](#).)

If her presence in these lists was really about including historically important visualizations,

## From our sponsors



Its the gift that keeps on giving!

SUBSCRIBE  
NOW! \$40 for 2 issues

Just in time for the holidays!  
Give the gift of Nightingale Magazine!

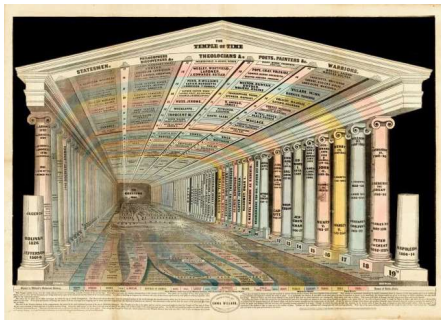


[SIGN UP FOR THE 'GALE](#)

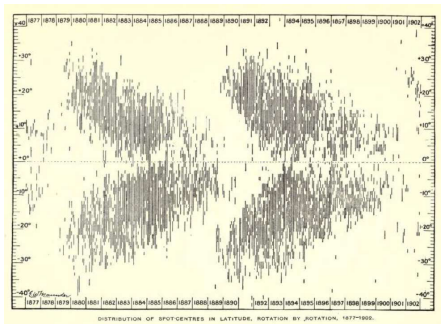
Get our monthly newsletter! It's a fun overview of our content and community

[Sign Up](#)





"Willard's Temple Of Time", Emma Willard, 1846 ([link](#))



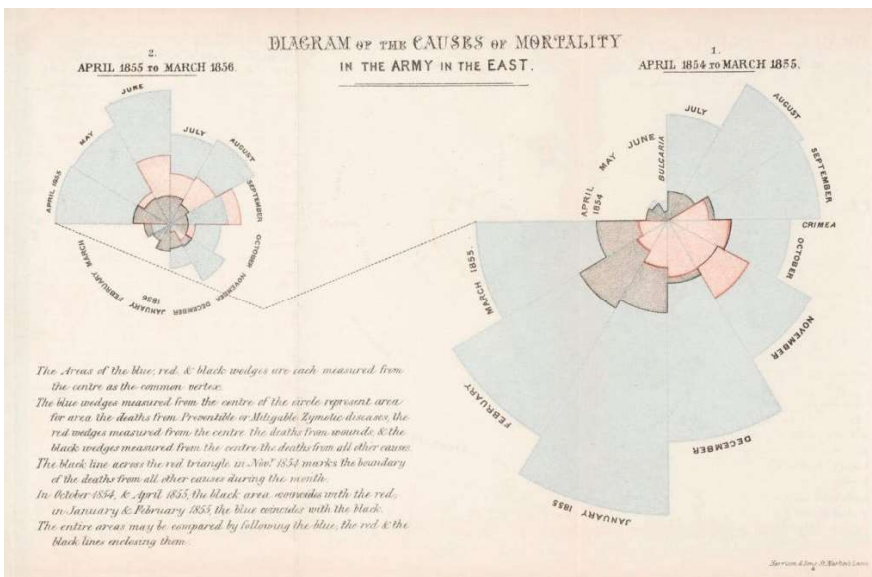
Distribution of sun-spots by latitude and rotation, Annie Munder, 1908, ([link](#))

we would see those same lists and books and classes include many more female visualizers throughout history. Who, you ask? Take [Emma Willard](#), a pioneering educator who revamped the way geography was taught by designing her own set of more accurate and contextual maps. While typical geography curricula used maps that focused on the thirteen original US colonies, thereby starting the discussion at the American Revolutionary War, Willard drew maps based on Native American territories (though they would likely be disputed by Native Americans) to focus the conversation on what was happening on the land before the invasion. They would also include [Florence Kelley](#), [Annie Maunder](#), [Marie Neurath](#), and [Mary Eleanor Spear](#)— just to name a few. Stay tuned for more profiles on these women to come soon in Nightingale and in the meantime, look up the work of [Catherine D'Ignazio & Lauren Klein](#).

So why is Nightingale the token?

I'm no historian, but I suspect it's in part because the accounting of viz history is rooted in Europeans and their wars: Nightingale's Rose chart on soldier deaths, Minard's alluvial diagram of Napoleon's march, Playfair's original pie chart of the domination of the Turkish empire, and so on. The scope of our field's history, like so much of the way history has been

framed, is Eurocentric in view and rooted in conquest.



"Diagram of the Causes of Mortality in the Army in the East", Florence Nightingale (1858)

If we wanted a fuller picture of the history of data visualization, we would need to include [North African cave paintings](#) dating back to 6000 BCE that tracked shifts in weather and migration. We would need to include the [Aztec calendar](#) that was our original timeline (well, time circle — just goes to show you that linearity is not universal).

These historically important visualizations go far beyond Europeans and wars but have been almost entirely omitted from historical discussions primarily conducted by cis white men about cis white men. (If you want to stop reading because of this phrase I challenge you to first ask yourself why and then to keep reading to see what you can learn from being uncomfortable.) When history is written to uphold the dominant power structure time and time again, it's no wonder we get visualization history that includes cis white men and the

## DVS Membership + PRINT!

Why not become a DVS General Member \*AND\* upgrade to receive a print subscription with your membership for \$25 (a \$15 discount). Patron Membership also includes print subscription to Nightingale Magazine.

[Learn more!](#)

token white female. (Just to be clear — white women and people of color have most certainly had their role in upholding various aspects of this dominant power structure, too — Florence Nightingale included. But given clear patterns of oppression, many would agree that everyone else is following the cis white male lead.)

Women have always had to fight for credibility and recognition in a male-dominated world.

Nor is this struggle specific to women exclusively. We know that the contributions of many non-male, non-cis, non-white people have been rendered invisible over the years. And while you'd think visualization (or really any major field like science, mathematics, or engineering) would get visibly more inclusive as time moved forward, such has not really been the case. In the modern era of data visualization, many point to Cleveland, Tufte, and Few — who all share a particular demographic in common — as forefathers. But how many of the foremothers referenced above could you name? Have you been taught the legacy of their impact on the field, or seen the countless examples of their work? It would be absurd to assume these women didn't exist — when really they fit an unfortunately all-too-common historical pattern that their work is simply not recognized by the dominant voices throughout history.

Even today, where the data visualization field is large and international, cis white men still dominate, through both hidden and overt ways. Men and women alike are probably able to identify some of the more overt ways, so let me tell you three stories about the more hidden ways that data visualization is still a man's world, despite the efforts of women (and some men) to change it.



Illustration by Jen Ray

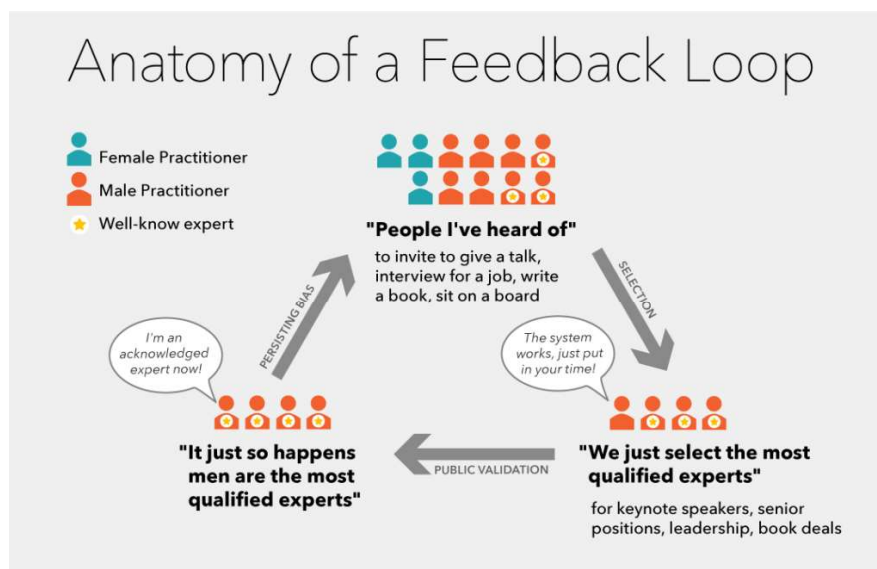
## Story One

At the end of 2014, I listened to [an episode](#) of a prominent data visualization podcast in which four cis white men, located in different parts of the world, in which they reviewed the cool visualization milestones and important moments for our field that year. *You know, providing recognition and credibility.* In their episode and in the show notes that accompanied it, they didn't feature a single woman. The kicker was that, at the end of the podcast, they lamented how hard it was for them to get noticed. All four of them have books, cited articles, talks, and/or widely-recognized visualizations. While we all may feel recognition is hard to come by, a group of fairly visible men lamenting they're not **more** visible, while directly contributing to the exclusion of women from the very narrative of our field, demonstrated a complete lack of self-awareness. In other words, if they thought they had it so bad, try being a woman (or person of color for that matter) in this field. It led me to write a [blog post](#) (not updated) that culls together a list of women in data visualization.

My post led to praise and anger on social media from men and women. Some women did not want to be pointed out as a woman. They were afraid of being tokenized, included just because they are a woman. Some men were defensive, claiming they just didn't know that many women in the field, and that intentionally identifying them would feel tokenizing or even stalker-y. It is a complex tension, isn't it?

Two years later, the same four men had another year-end [episode](#) and I remarked that it might be nice to have a woman on the show to add to the voice of authority about what was Important That Year. The same tensions arose around tokenism. The thing is, women (all people) want a voice at the mic because of their good work, not their gender. People in a place of privilege should highlight females because their own understanding of what's happening in the field is broad enough to include the incredible work being done by women. While this piece isn't meant to be a primer on how to avoid tokenism, **engaging equally (in time spent, attention paid, payment rendered and respect extended) with the full body of work that countless women have**

produced is a pretty solid way to avoid it.



Anatomy of a Feedback Loop cycle illustration by Elijah Meeks and Amy Cesal

## Story Two

In between those episodes, I started my own podcast in part to grow my own credibility and in part to highlight others. As I was logging into Skype to record our 10th episode, I heard my co-host and the episode's guest (two cis white men) chatting. My audio had connected before my video. And I overheard them talking about my appearance and discussing whether they found me attractive. Then I heard one of them say "Oh, looks like she's joining" and they abruptly stopped the conversation. I didn't really know how to handle the situation. I did what a lot of women do on a regular basis: Smile and act like nothing happened. Gloss over the injustice and try to be a professional. Grin and bear it.

The commentary, however unintentionally harmful, undermined my credibility and damages the ability for everyone involved to provide the best of ourselves to the world. It makes me and every woman who has witnessed something similar have to fight twice as hard to be heard.

## Story Three

Just this year, the aforementioned podcast guest from episode 10 contacted me with an invitation to be a featured speaker (for free) at a conference he regularly organizes. I realized this was an opportunity to respectfully reply that my participation was contingent on the conference abiding by an inclusion rider.

Popularized by Frances McDormand in her Academy Award acceptance speech in 2018, an inclusion rider is a stipulation that requires diversity. The specifics are usually negotiated, but in my case, I asked that the other featured speakers proportionately reflect United States demographics in terms of race, gender, disability, and LGBTQIA+ status. An inclusion rider is one way that people in places of privilege can fight together with those who have less privilege for an equal spot at the podium.

His response to my request was that he would be uncomfortable finding out those demographics, particularly the latter two. I pointed him to Frances McDormand and the group she was working with and their STACKS of resources on how to go about fulfilling an inclusion rider — all of which was easily Google-able. And I didn't hear from him again. No featured speaker slot means no boost to my recognition and credibility.





Illustration by Jen Ray

My three stories are not just mine. These happen in some iteration to women all across our field (and in every other field, too). Many women share my frustrations with getting equal play in a male-dominated field. Some defensive men might point to the well-deserved rise in popularity and recognition for individual women. Nadieh Bremer, Giorgia Lupi, Stefanie Posavec, Amy Cesal, Mona Chalabi, and Michelle Rial are just some examples of women who have been making waves lately. Their work has an artistic aesthetic with a human touch — often hand-drawn, made with markers, colored pencils, or play-doh. Their work is awesome and deserving of praise but that playfulness is in stark contrast to the spartan, tech-oriented, traditional aesthetic we commonly see from impact-focused thought leaders.

In other words, I just don't see many legions of cis white men following their lead the way they line up behind other cis white men in our field. And when cis white men do, they are freely allowed to bounce back and forth between the artistic and spartan aesthetics, whereas these women have been somewhat pigeon-holed and not socially graced with the same freedom. Moreover, I would put money on the notion that these women have stories similar to mine.

The problem is most definitely not that there's only one group capable of being high-quality data visualizers. I have heard, in the darker corners of the internet, commentary along the lines of "women don't make visualizations up to the standards" which is the data-based version of "women just aren't funny." This just begs the question of — by whose standards? Literally, who wrote them? This field is young enough that we can actually find out. Tufte, Brinton, Few and others articulated their standards pretty clearly actually. They have one particular set of characteristics in common. Standards that don't specifically include other viewpoints perpetuate the historical status quo.

The problem with stories one and three is that they limit the scope of the field. If you only see/recognize work that resembles your own, or by people who resemble you, then your definition of the data visualization community is inherently tiny. Similarly, if you only look at the people in your Twitter feed, you are missing out on a lot of cool ideas and awesome visualizations made by people who happened to be women, of color, with a disability, and/or LGBTQIA+.

So when leaders in our field produce newsletters that predominantly feature the work and thoughts of white cis men, it perpetuates the notion that their views are the ones that matter AND that there aren't others outside that demographic at all. Stephen Few primarily referred to men's thoughts and ideas in his newsletter, often referencing something one of them had said in the comments to a previous edition of his newsletter, creating a cis white male echo chamber. Why did he cite these men's comments? Because men were the vocal majority in his comment thread, a place widely seen as combative and hostile (reference my earlier discussion of men and war). While some in the data visualization world find that trashing Few or Tufte on Twitter will earn you a lot of likes, the fact of their influence on this community is unavoidable.

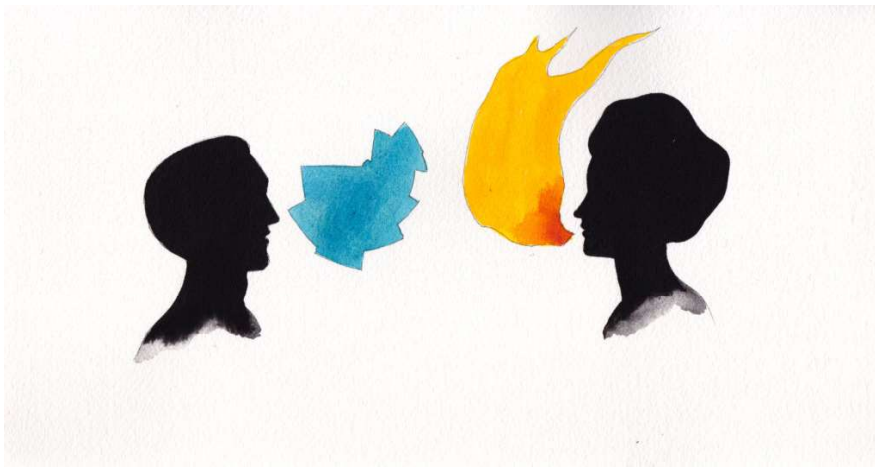


Illustration by Jen Ray

## What does it take to get beyond Nightingale?

This is the section of the article where I conclude with positivity and solutions. But I am going to depart from the long tradition of women asking (begging?) men to change their behaviors. I am not going to provide a bulleted list of concrete solutions and steps to take. Why not? Because those answers are already out there. They have been said (screamed?) for generations. **You just have to pay attention.** Women work hard for recognition (and non-white women work even harder).

While there are more sources of credibility and recognition, who we cite and who we invite as speakers and who we showcase on our podcasts are important contributing factors. There's plenty of history-making to go around.

### Author profile

#### STEPHANIE EVERGREEN



Website

Dr. Stephanie Evergreen is an internationally-recognized data visualization and design expert. She has trained future data nerds worldwide through keynote presentations and workshops including Facebook, Time, Adobe, Verizon, and the United Nations. She writes a popular blog on data presentation. Her books, *Effective Data Visualization* and *Presenting Data Effectively*, both hit #1 on Amazon bestseller lists.

### Related entries

This author does not have any more posts.

### CATEGORIES:

DATA HUMANISM

HISTORIC DATAVIZ

TAGS: [data feminism](#), [data humanism](#), [data visualization](#), [equity](#), [historic dataviz](#), [women in tech](#)

[← PREVIOUS ARTICLE](#)

**Visualizing Data for Audiences in Low & Middle-Income Countries**

[NEXT ARTICLE →](#)

**Intricate & Visually Exciting: An Interview with Nadieh Bremer**

Related articles

Menu

Reading: Asia's data scene deserves greater attention. That's why we are starting a movement.

# Nightingale

Journal of the  
Data Visualization Society



Home

Topics in DataViz

How To

Community

About

SUBSCRIBE

## TOPICS IN DATAVIZ

Asia's data scene deserves greater attention. That's why we are starting a movement.

Loh Pei Ying • August 25, 2021



Various illustration and data examples from Kontinentalist's stories.

A data-driven studio telling stories about Asia, Kontinentalist uses data storytelling to bridge the gap between research and the public, bringing Asia to the forefront of global conversations. It is a part of Potato Productions, a diverse group of companies combining technology and creativity to achieve social good.

### Why does Kontinentalist exist?

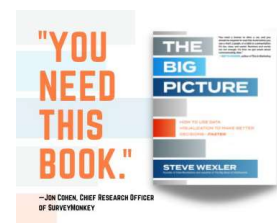
Kontinentalist didn't set out to be data-driven. Our director Lee Han Shih (a retired journalist with 30 years of experience) and I shared a common passion: we wanted to tell stories about Asia on its own terms. Having learned about Eurocentrism and Orientalism as a history major, I wanted to help unpack misconceptions and myths about our region's past and present.

We founded Kontinentalist with a simple purpose: to tell the story of changes in Asia, make quality information accessible, and keep communities updated on what's happening around them.

### Stumbling into data visualization

Our transition to data visualization started with China's Belt and Road Initiative. Announced in 2013, the initiative seeks to reinvigorate historic trade routes, cultivating multilateral economic cooperation and connectivity with infrastructural developments. A juggernaut of a project set to roll through three continents, it was massive and controversial—but there was so little reliable information on it out there. We felt this needed to change.

## From our sponsors



Its the gift that keeps on giving!

SUBSCRIBE  
NOW! \$40 for 2  
issues

Just in time for  
the holidays!  
Give the gift of  
Nightingale  
Magazine!



SIGN UP FOR THE 'GALE

Get our monthly newsletter! It's a fun overview of our content and community

Sign Up

Having worked in museums before this, I knew next to nothing about tech, data, and design. To tell a data-driven story about the Belt and Road Initiative, I would first have to teach myself how to collect data and tell stories with it.

So, the team signed up for lessons on [datajournalism.com](https://datajournalism.com). I remember feeling mind-blown by *Reuters*'s "Connected China" project and *South China Morning Post*'s multimedia visual explainers. Needless to say, we were hooked. We bought data visualization books and observed the best in the business: the *New York Times*, *The Pudding*, *Washington Post*, and many others.

It was an intense process of self-learning and emulation. We practised our newfound skills with our first map-scroll piece, [Understanding the Belt and Road](#)—and the rest is history.



Infrastructural projects along the Belt and Road as of November 2018. The data on this map was produced by Kontinentalist, Singapore. Full story: <https://bri.kontinentalist.com>

## Finding a community here in Asia

The deeper we got into data visualization, the more we felt challenged by how out of reach the industry's resources and communities often were. When we "discovered" the data visualization scene, we wanted to join conferences, meet new people, and learn through on-site courses... but so much of this stuff only happened in the Western world. Even when events were digital, they were held in unearthly hours for our time zones.

This made our learning journey quite lonely for a long while. But it also pushed us to look for companions nearby! We found the [Singapore Data Viz meetup group](#) and started befriending folks there, and this helped us feel a sense of belonging. In the past few years, the data visualization scene in Asia-Pacific has been quietly growing. We've found new friends such as [Punch Up](#) in Thailand, [Open Development Mekong](#), [Synthesis](#), [Nugit](#), [Gurman Bhatia](#), and many others.

But we wanted more for Asia's data visualization community. What else lay out there?

## The Outlier 2021 conference

When the Data Visualization Society (DVS) announced their plans for their very first conference, "Outlier" (with sessions we could attend from our time zone!), we were stoked. We'd been members for a while by then, and we were—and still are—always grateful that such a wonderful ground-up organisation existed for the community.

We wanted to contribute to the collaborative, open, and sharing culture of DVS. When it called for informal "unconference" sessions, we saw it as our chance to add something special, unique, and fun. As a shot in the dark, we decided to list "Are you an Asian snack?" as a theme, hoping that the universal love for food would



## DVS Membership + PRINT!

Why not become a DVS General Member \*AND\* upgrade to receive a print subscription with your membership for \$25 (a \$15 discount). Patron Membership also includes print subscription to Nightingale Magazine.

[Learn more!](#)



help us find some common ground among conference attendees.

Frankly, we were worried that no one would show up. As it turned out, we were dead wrong—a full 20 odd people showed up, snacks in tow and eager for some good conversation. The level of participation was mind-blowing. Although we had planned for a casual and fun session, the conversation started to get real serious.



Participants of the "Are you an Asian snack?" unconference session at Outlier 2021.

We spent more than an hour talking about the challenges faced by Asian dataviz practitioners and the scene in Asia. We talked about language barriers, cultural misunderstandings, and gaps. Trying to fit into an editorial world and style centred on the West. Our hesitation to promote ourselves and our work because it felt so un-Asian—or because we didn't feel we had the right to do so.

Many of us didn't feel like we stood a chance at international awards in our own field. Could we start our own awards? What would that look like? Time passed, we overran our slot, and still it wasn't enough. We left feeling like we'd barely scratched the surface—but we'd also found a community, and we wanted to do more.

### What's happening in the Asian dataviz community?

After hosting that unconference session, we felt we needed to continue the work. We went back to the drawing board, had more deep discussions with colleagues and friends, and identified some main challenges to tackle.

Asia has a big data gap. There is plenty of data about Asia, but most of it is dispersed. Countless non-governmental and nonprofit organisations in Asia invest their money into research and data programs each year, but where does this data go, and how accessible is it? Who reads it beyond its niche audience of agents and researchers?

Most Asian countries rank low on the Global Open Data Index; there are plenty of reasons for this. Between a lack of publicly available official data and poor freedom of information legislation across the region, it's no surprise that data visualization work is not as well-established here—there is simply less open data to work with.

Asia's language diversity is one of its greatest strengths, but it's also often a stumbling block to data work. Most resources on working with data are published in English, with only a small number in other major languages. Many Asian dataviz practitioners, especially those doing work for smaller communities, cannot access this material. Their work is also less likely to gain recognition without sensitive, high-quality translation—which is often expensive and time consuming.

These challenges extend even beyond the data itself. How might we bring an Asian perspective to frameworks, methods, and thinking in data visualization, so often curated through a Western lens? As a history enthusiast, I'm excited by the rich history of Western data visualizations—but where is this history in Asia? There are some amazing examples of data visualization from Japan (shout out to [RJ Andrews](#)), but we can't help but wonder if there's more out there, waiting to be discovered.

Geographical borders, for example, did not exist as a concept in pre-colonial Southeast Asia—are there ways to visualise the region's histories without them? How about cultures with their own units of measurement or unique ways of telling time? What Asian motifs or forms lend themselves to data visualization?

We have witnessed this ourselves: data visualization is a creative and experimental medium, and those characteristics have allowed many Asian practitioners to add their own cultural elements to their work. We've seen how traditional Asian art forms can be used for data visualizations, such as [henna art](#), [embroidery patterns](#), [calligraphy](#), and more.

## Our hopes and dreams for Asia's dataviz scene

Economic observers often talk about Asia's potential to shape the world. We agree, and not just because of our region's rising affluence. Asia's diversity will be—has already been—a major game changer in so many fields. Our cultures hold new perspectives, creative ideas, and fresh information—and the world is paying attention.

As data practitioners in Asia, we have front-row tickets to this incredible transformation—even as we each bring our gifts to the world stage. Kontinentalist is a small entity, and our reach is modest, but we're working hard with our fellow dataviz practitioners here to address the challenges we all face.

We're starting by partnering with as many causes as we can to bridge the gap between their data and the public. We want the NGOs, non-profits, and research groups here in Asia to realise the full potential of their data. We can all do good work with this, even if official government data remains hard to come by.

We've been chipping away at this for two years now. Our [latest collaboration](#) with [Médecins Sans Frontières/Doctors without Borders on the Hepatitis C epidemic](#), for example, is meant to help uplift causes that we believe deserve more attention.



An illustrated map depicting the physical distance Rohingya refugees have to make to get healthcare in the camp. This map is part of a larger story we did on inequality of access to Hepatitis C care in Asia, in partnership with Médecins Sans Frontières/Doctors Without Borders.

Championing Asia's cultures through data is also a big part of the work that we do. People often say it's difficult to tell stories without good data, but many don't realize that they're only looking at quantitative data or assume that data must be statistically tested or peer-reviewed for it to be usable.

At Kontinentalist, we challenge this by developing stories that reveal the data all around us. Data is in our languages and cultures; we just need to take a closer look. We use our platform to spotlight aspects of our cultures that we love and celebrate. It's also an incredible opportunity to be at the forefront of this change and to see our own cultures get the representation they deserve—delivered with proper context, nuance, and visuals that capture their spirit.

Since Outlier 2021, we've been wondering how we can contribute to this newfound community and help keep its momentum going. Our new friends had so many personal tales to tell, and so many lessons to share, that we decided to dedicate our ["Community" series](#)—an ongoing series of interviews—to Asia's data practitioners and enthusiasts. We want to amplify their work and create an affirming space where members of our community can connect with one another.

It's a small effort right now, but we hope to start a movement with it. Asian data practitioners should be proud of their identity and their work. We hope our platform encourages them and allows them to share their challenges. As the world comes to recognise the need for diversity and inclusion, these voices have never been more important.





This is only the beginning. As Asians, we often joke that we've been brought up to achieve good grades and overachieve academically. But this shows up in the data landscape here, too. Most data visualization courses target business analytics, dashboards, or Big Data. We hope to blur the lines between business, science, and art, and to show that there is more to the industry than meets the eye. We have plans to launch data visualization workshops in Singapore for newcomers by the end of the year, with the aim of connecting them with experienced professionals so they find a community that keeps them going. We hope that with this small but close community, they may push the boundaries of imagination and create unexpected, beautiful data stories.

Asia's dataviz scene is coming into its own. It may seem nascent, but this just means there's greater flexibility and room for change. We want to be with it as it grows and succeeds. Many people have reached out to affirm our efforts and let us know we're on the right track; we hope you feel the same. Join us in growing this movement—or even start your own!

#### Author profile



#### LOH PEI YING

Website

- ☐ Pei Ying wears many hats in Kontinentalist. She leads the company in achieving its overall business and editorial goals, making strategic business development plans, and managing partnerships. Her background and passion for history is the driving force behind many of her stories, which delve into cultural and historical contexts. In her free time, she is likely tending to her veggie garden, cooking, or cuddling her two fat cats.
- ☐
- ☐

#### Related entries

This author does not have any more posts.

#### CATEGORIES:

[ALL STORIES](#)
[COMMUNITY](#)
[DATA JOURNALISM](#)
[EQUITY IN TECH](#)

TAGS: [asia](#), [community](#), [data journalism](#), [data storytelling](#), [data visualization](#)

[← PREVIOUS ARTICLE](#)

**Making Sense of Satellite Data, An Open Source Workflow: Accessing Data**

[NEXT ARTICLE →](#)

**REVIEW: Data Vis Book Club**

Related articles

# Reading Unit 4: Graphical Distortion

(Estimated Reading Time: 1.5 hours,  
Complete Reading by 8am 02/21/2022)

**UO41:** Through readings, group discussion and reflection, students will evaluate misleading graphical summaries of data with an equity lens to explore how graphical summaries can exacerbate structural oppression. (CO 12)

**UO42:** Through readings, group discussion and reflection, students will identify best practices for promoting equity through appropriate data visualization techniques. (CO 12)

1. Crick, S. (December 14 2019). Be Mindful of Misrepresenting Data in Visualization. <https://shelleycrick.com/be-mindful-of-misrepresenting-data-in-visualization/> (accessed 2022-01-08).
2. Schwabish, J., & Feng, A. (2021). Do No Harm Guide: Applying Equity Awareness in Data Visualization. <https://www.urban.org/research/publication/do-no-harm-guide-applying-equity-awareness-data-visualization>.

# Reading Unit 4:

## Discussion Questions

1. In the examples given in the first reading, how do the graphical mistakes made reflect the implicit biases of the creators? How does the author equate the creation of misleading graphics to power and privilege?
2. What techniques do the researchers at The Urban Institute recommend to create equity within and through data visualization? Do you agree with their assessments?

# Reading Unit 4:

## Reflection Questions

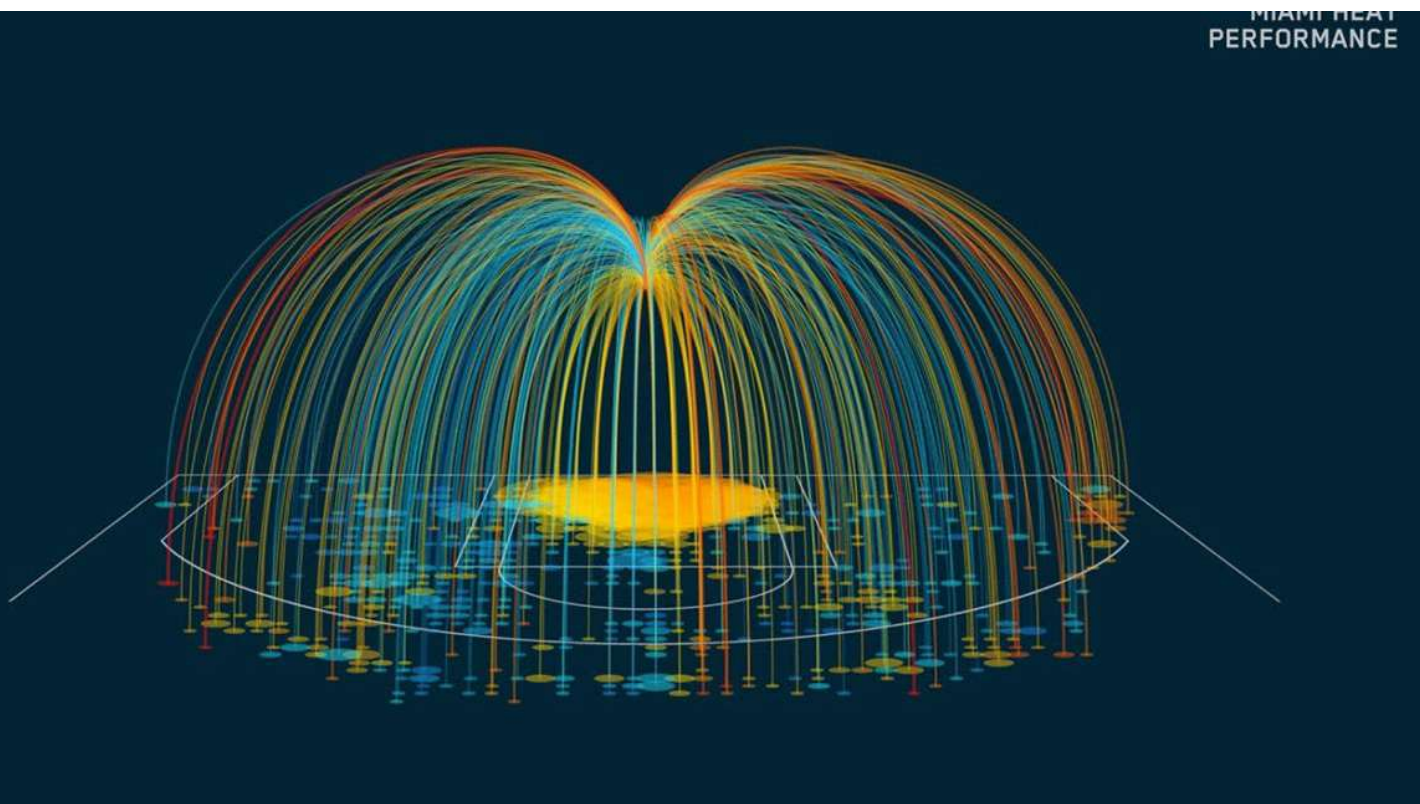
1. Thinking in terms of your major and your future career, which types of data misrepresentations in statistical graphics do you think you will be most likely to encounter?
2. What steps can you take in your future career to ensure that any statistical graphics you create contribute to a more equitable society?



shelleycrick

## Be Mindful of Misrepresenting Data in Visualization

Posted by Shelley Crick | Dec 14, 2019 | Data Visualization | 0 | □ □ □ □ □



It can be challenging to find creative and effective ways to present data to an audience. Some folks who participate in research and study data enjoy chewing on pages and tables full of numbers, but most general audiences do not. They are looking to glean and retain the most critical pieces of information as fast as they can.

When you're choosing how to share your data in the most effective way possible, it's vital that your summarized version still tells an accurate and complete story.

Data visualization provides a simple way to **see and understand trends** and patterns in data and can make it possible to analyze vast amounts of information at once. Since our eyes are drawn to colors and patterns, data visualization helps keep an audience focused on the message. It's storytelling with a purpose.

With the freedom and flexibility that comes with communicating data through visuals, there are guidelines to follow, or you could end up causing more harm than good. Unintentionally—or even intentionally—misrepresenting or manipulating data jeopardizes the credibility of your visualization, and it is unethical to leave out pieces of data because it disproves your point or doesn't jive with your narrative.

And yet, manipulation of data in visuals is a common practice.

## What is Data Visualization

Data visualization is a hybrid form of art and science. Here, how the data is presented is as important as the data itself because **an image invites the imagination to participate** in the conversation. Typically in the form of charts, tables, graphs, maps, infographics, and dashboards, data visualization provides a way to organize numerical data—sometimes large amounts of complex statistical data—into digestible visual formats. This helps a viewer better understand information and to see patterns in data.

Some of the **common uses** for data and information visualization are:

- **Presentations for understanding:** Some ideas are too awkward to express in words, but translate well in graphic formats. Things like maps or city plans.
- **Explorative analysis:** Using charts and graphs to uncover patterns that require further research.
- **Confirmation analysis:** Helps to confirm an understanding of data and to explore relationships and correlations further.

Visualizing big data can reveal hidden patterns and uncover lesser-known relationships to improve decision making. But with so much data coming from so many sources, it can be **challenging to provide accurate, real-time visualization** to explore big data. The nature of big data makes accuracy and completeness difficult to identify and track, which increases the risk of false discoveries, or skewed findings.

In the fields of science and mathematics, data can be hard to communicate to an audience that's unfamiliar with the language. Although historically very limited, graphics and images have become essential translators of scientific matters to the general public. The problem is that science communicators are typically **not trained in visual communication** and unintentional errors and misrepresentation can happen. Pair that with the knowledge that the tools for visual storytelling change every day, it makes it more difficult for researchers to **keep up** with the challenges of scientific data visualization.

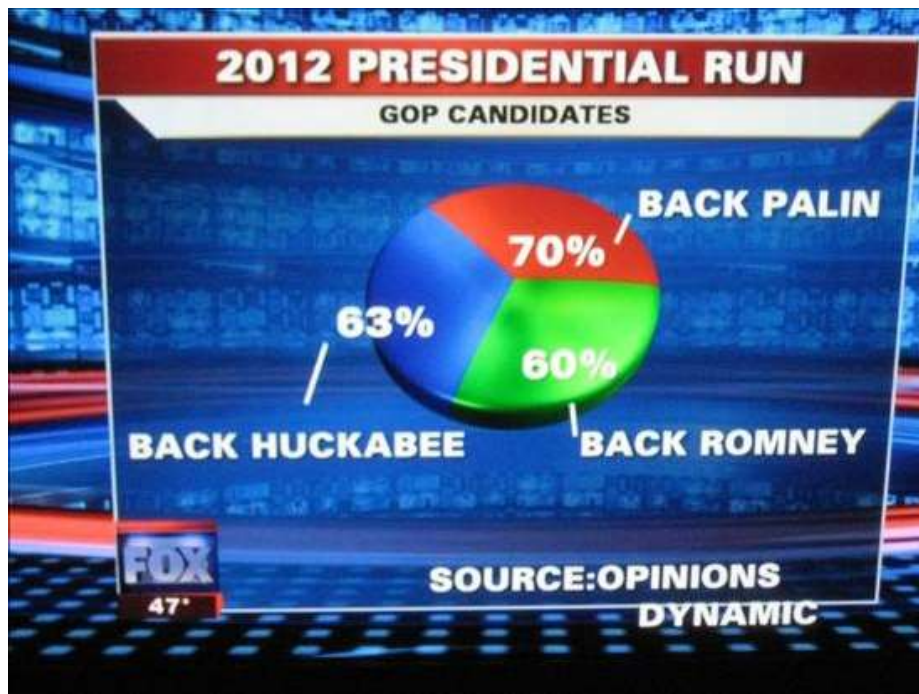
While it's true that the numbers don't lie, they can be used to **mislead and tell half-truths**. Often, it's assumed that companies or stakeholders intentionally misuse statistics because they have something to gain from distorting the truth. Sometimes though, it's design mistakes.

## How Can Visuals Be Manipulated?

The global accessibility of data means it can be harder to find the right information to tell a particular story, so a visual storyteller has to be very careful where they get their data and how they share it.

Assuming you've used a reliable data source, some of the ways perfectly sound and useful data can be manipulated are by;

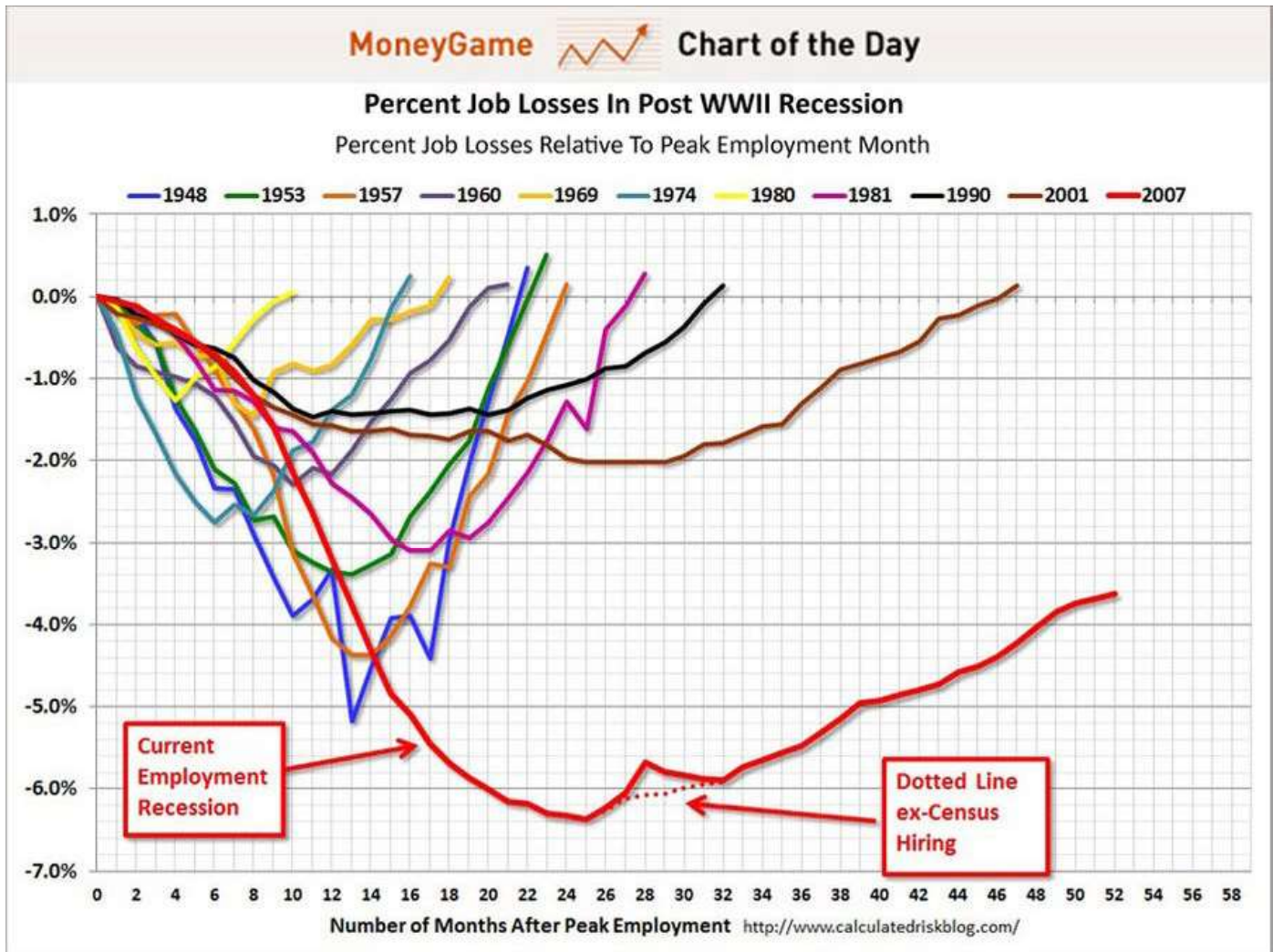
- **Misrepresenting data:** Beware of **mislabeling data** or using inaccurate descriptions for a visual. Or even, using the wrong chart. And when presenting data that shows a correlation, don't imply the relationship. Don't be afraid to add context because that may help the viewer understand what they're seeing.



FOX 2012 Presidential Run pie chart has problems. A bar chart would have been a more logical display.

- **Using too much data:** Your target audience needs to be able to look at your graphic and find what they are looking for without clutter. When a graphic is too busy, it can be hard to find the main points. So keep it simple and narrow the focus of each graph.

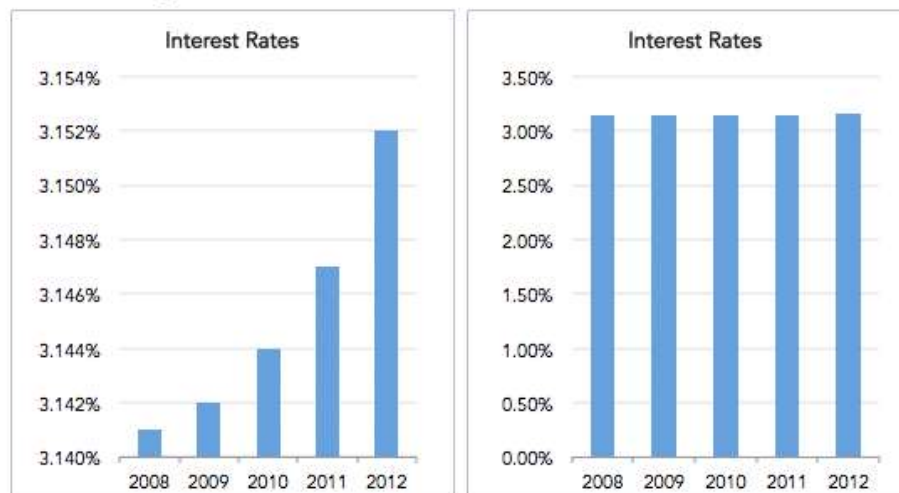




### Job Loss Post WWII Recession

- **Distorting data:** Again with using the wrong chart type, not charting to scale, omitting data which may hide trends while creating trends that don't exist, and truncating the y-axis are all ways that data can become distorted. These distortions make it hard to compare data points.

### Same Data, Different Y-Axis



## Truncated y-axis

Graphics created electronically carry more authority than images drawn on **paper; people expect computers to be right** where hand-drawn images are expected to leave room for a bit of error.

There has also been concern among the science industry that computer-generated graphs and graphics could be tricking the public—and even the scientists who use them—into believing that theories or forecasts were proven facts because of their finished presentation. Speakers at the Conference of the American Association for the Advancement of Science **noted in 1994** that in “an image-hungry world, a computer forecast of patterns of air pollution was more effective in influencing policymakers and politicians than dry tables of numbers and charts.” This caused the speakers to worry about the quality control of graphs or graphics, because the finished products were professional enough to pass for the truth, without the same intense peer review as scientific papers.

## Examples of Inaccurate Visuals

The ability to use graphic software to create visuals gives scientists a way to explore their data in methods that may have previously been cost-prohibitive. Now they can create graphics themselves! Cost-effective it may be, but cutting out the trained graphic designer can also lower the standards of their completed work.

On the flip side, a cartographer drawing a map uses their years of science experience and training to create the most accurate map they can. A graphic designer without cartography, archaeology, or GIS training creating a map may produce a better quality or cleaner version. Still, without that training, it's more likely that the designer creates a beautiful but **misleading map**.

Spend any time on the internet and you will eventually come across an inaccurate graphic with potentially misrepresented data. Here are a few examples that have made the rounds.

### 1. Prison/Princeton

This graphic created by PublicAdministration.net made the viral circuit a few years ago **but is incorrect**. Believe it or not, the problem with this otherwise beautiful graphic is that the numbers for Princeton and prison are flipped. The annual tuition cost at Princeton in 2010 was \$48,500, and the cost of incarceration in New Jersey in 2009 was \$38,700. They've got it backward.

In this example, someone had a specific story they want to tell and manipulated the data to tell that story. “Gasp! We spend more on prisoners than an Ivy League college!” was the viral response. If that were true, this would be the wrong chart to prove it.



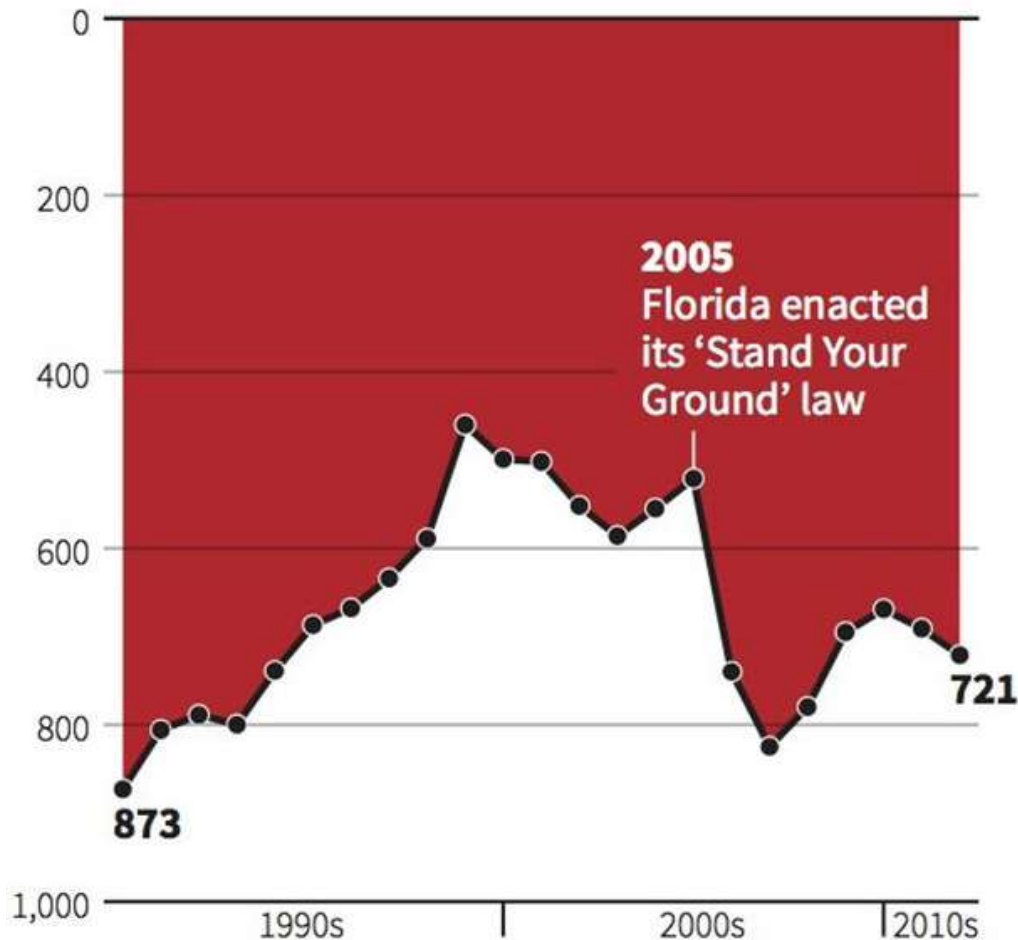
Prison/Princeton

## 2. Gun deaths in Florida

This Reuters chart is trying to show how the number of gun deaths in Florida has changed since the 'Stand Your Ground' law was enacted in 2005. The chart is actually **upside down**. The x-axis starts at 0 and goes down to 800, making it look like the number of gun-related deaths has declined when actually, it's gone up.

# Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

Florida gun deaths

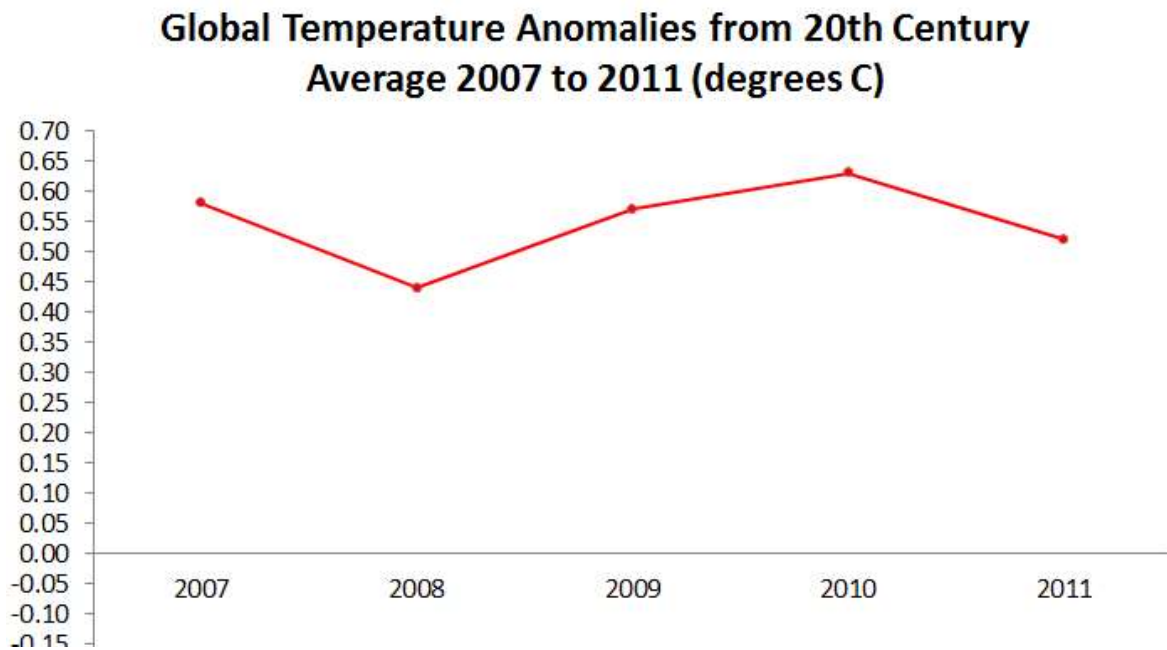
Why might someone present a chart this way? In all probability, this was the graphic designer taking a bit of a creative license to allow for the dramatic red background and was not designed with the intent to deceive. But again, it might depend on the stakeholders behind the chart and what message they're trying to send. If they're supporting the law, they want it to look like the number of deaths has gone down—so they point the chart that direction.

## 3. Global cooling

It is not uncommon for stakeholders or organizations to cherry-pick data to support otherwise unsubstantiated claims.

The average temperature of our planet naturally varies a little from year to year. Climate change naysayers take advantage of this reality by showing temperature data in small clusters of years instead of a larger span, like decades. Filter out the natural pulsing of the planets, and an overall rise in temperature is apparent.

This line graph, which tries to communicate that there isn't a climate problem and that temperature just varies a bit, leaves out some critical information. Primarily, those decades of climate variances we talked about.

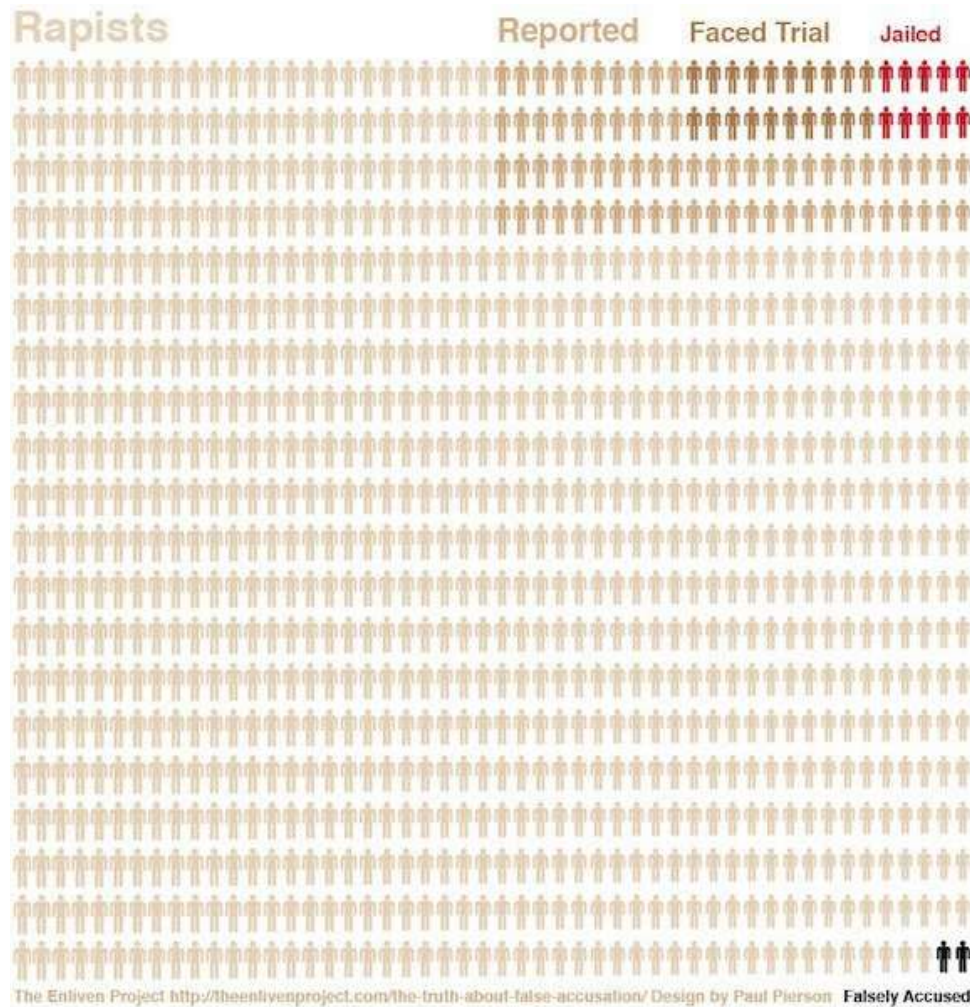


Global cooling

#### 4. The truth about false accusation

Sarah Beaulieu's viral infographic from 2012 was intended to drive home the point that in spite of what the media might suggest, false rape accusations are extremely rare. It does make that point; however, the rest of the graphic is inaccurate.





### The truth about false accusation

There is a difference between “false reporting,” claiming an assault has happened; and “falsely accusing,” which is pointing out and identifying a perpetrator. It’s important not to confuse the two terms. Research shows that **2-8 percent** of reported rapes are false, but the number of *accusations* is even smaller.

This graphic assumes one rape per rapist, but at the time it was created, the **average rapist had six victims**. By using the label “rapists” instead of “rapes,” this chart might give the impression that every-other person you meet is a rapist when in reality, it’s an estimated 6 percent of men, repeating their assaults.

Also, the Rape, Abuse and Incest National Network (RAINN) estimates **3 out of 4 sexual assaults** go unreported. This chart overestimates that number.

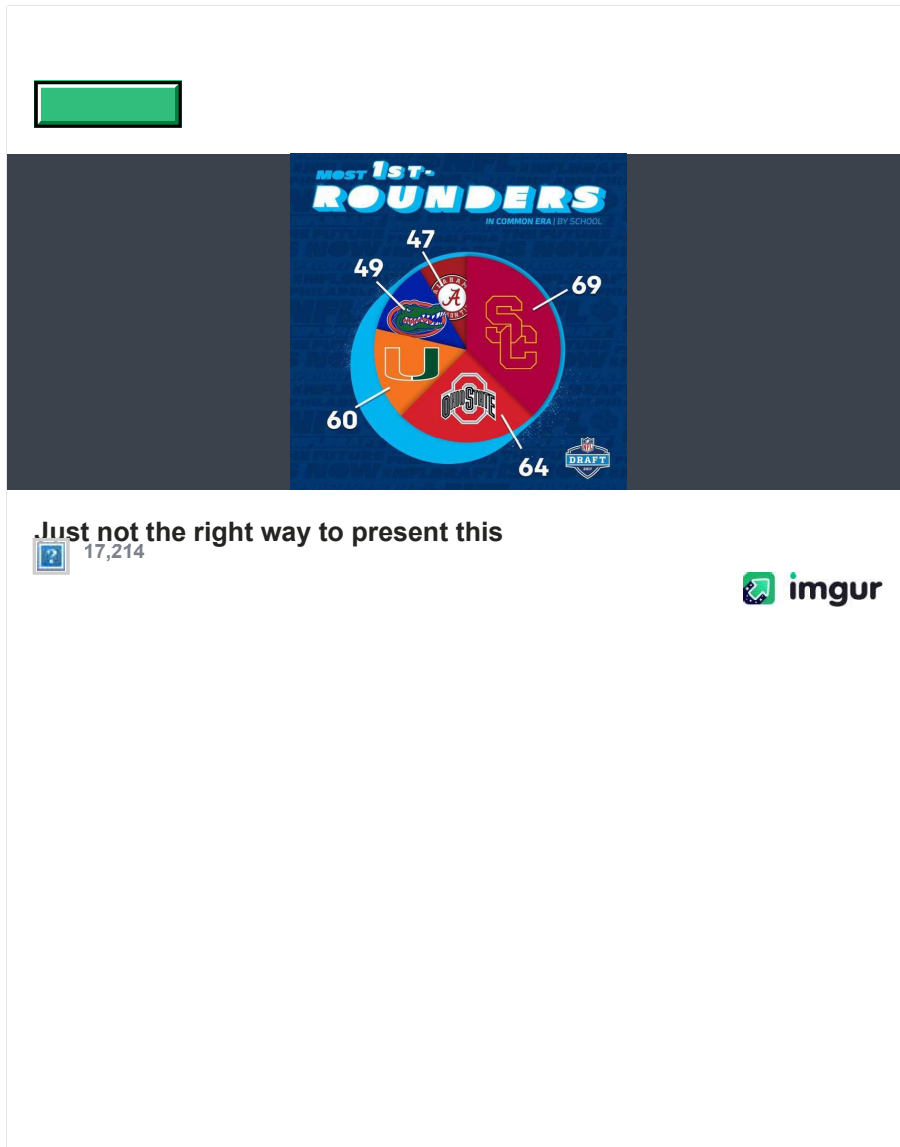
Why? Is it to point out how rare it is for an assault victim to report? Or is it to raise the conversation? Although it is correct that most sexual assaults go unreported and the chances of being falsely accused are small, the data in this graphic is telling an inaccurate story.

## 5. 2017 NRL Draft

Sometimes a chart is incorrect because it’s just the wrong format for the information provided. This chart



from the NFL draft is a good example.



As a pie chart is used to show percentages, the sections should add up to 100 percent. This percentage chart totals at 289. Based on the size of the parts, we're to believe that 60 is less than half of 69, and that USC is making a haul.

The thought is that the designer wanted to make all of the logos visible in their section, but the chart grossly misrepresents the truth. It's a great candidate for a bar chart.

## 6. The Big Cheese

Speaking of pie charts, this pizza version looks cool and means to tell a story about the executives who've served as CEO at Papa John's, but a viewer would never get that without reading every bit of text. The sections are unbalanced and this information would be much better presented as a timeline. Points for trying to be witty and topical, but it's also not the right format for the information.



The Big Cheese

## 7. Summer Olympic Medal Count

This graphic clipped from an edition of the *Erickson Times* shows a deceiving chart with its iconography not being to scale. Sure, there are more medals to show the U.S.A. awarded the highest medal count, but the number of medals shown in no way corresponds to the number of medals awarded.

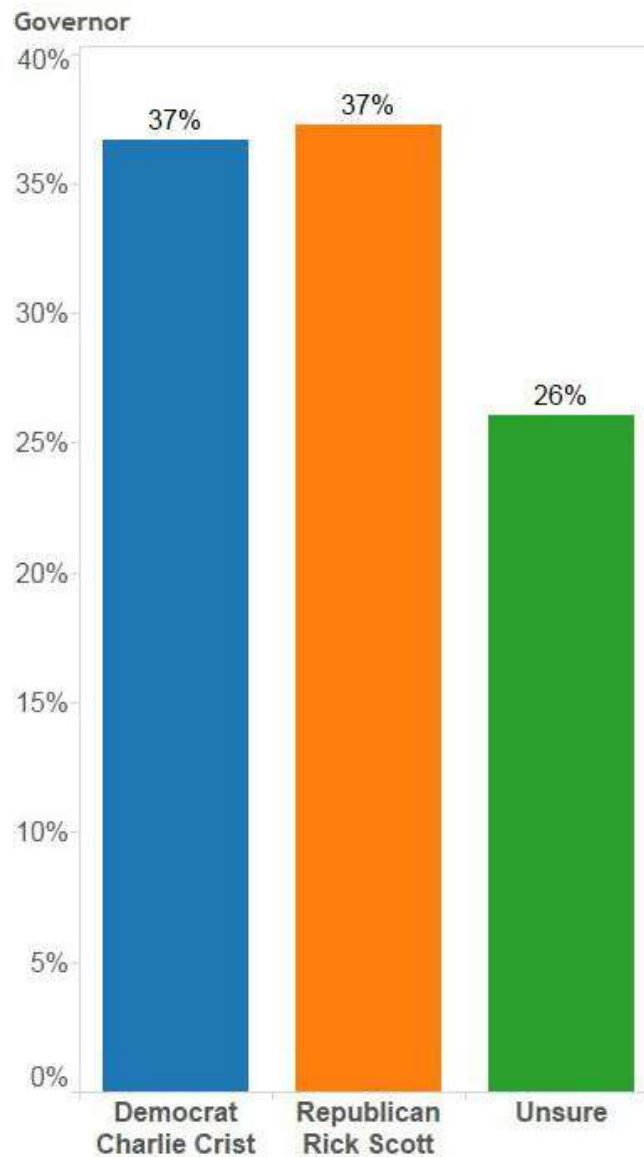
We're expected to believe that the two medals pictured for Germany represent their 499 medals. Then it instantly contradicts that assumption with three medals pictured for France's 523. And six medal icons represent 1,975 medals for the U.S.A. There is very little relationship between the images and the actual number of awarded medals. So why bother?



Olympic medals

## 8. Governor's Race

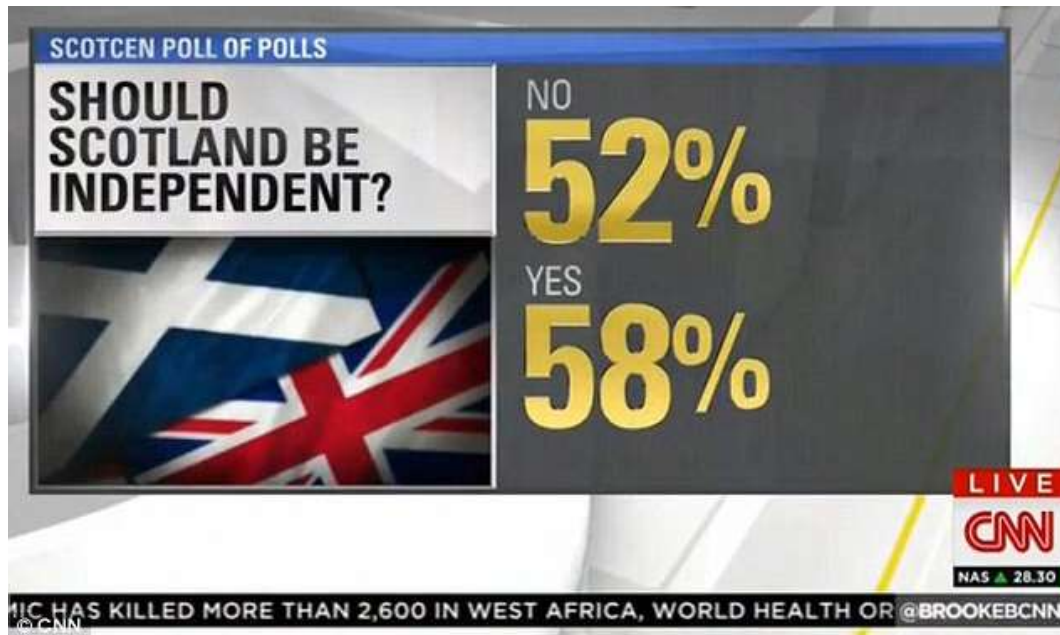
And then there was that very exciting Governor's race when one candidate's 37 percent was visibly larger than another candidate's 37 percent. There's no reason to go into what the problem is with this chart. However, since we're working with percentages and we add up to 100 percent (as is accurate), this might have been an ideal place for a pie chart.



Governor's race

## 9. Scotland gave 110%

Equally exciting was the incident when CNN gave us voting results as a data visualization in the form of simple numbers...and they were wrong.



CNN Scotland vote

The graphic showing the results of the voting for Scotland's independence was 52 percent 'no' and 58 percent 'yes,' for a total of 110 percent—a worrying result for a countrywide vote.

Not only was the data wrong, but it was wrong in the wrong direction. The actual vote was 52 percent 'no' and 48 percent 'yes.'

This example goes to show that it's not just bar and pie charts and infographics that can get it wrong and present inaccurate data; sometimes it's as simple as not checking your numbers.

## Reminder about the ethics of using your power wisely

Seeing is believing, and **people want to believe what they see**. Seeing data solidifies it in the mind's eye and makes it more real. Literate communication presented in tables or a line of text is boring, so regardless of how abstract or complex an immersive data presentation is, we are less skeptical of data presented in an immersive visual format. Data visualizations are so persuasive because they are both literate and immersive forms of communication, and we like them. So the information you present in visual formats needs to be accurate.

As data sets become more complex, displays should become **increasingly informative**, illuminating relationships that would be inaccessible from tables or summary statistics.

The multitude of ways data visualization can go off the rails is not a good argument for not using them. They can make complex information more transparent and are easily accessible. They just need to be accurate.

So before creating or sharing a data graphic, take a thorough look at it. Do a bit of mental math. Evaluate where the graphic—and even the data—came from. Is it from a reliable, well-regarded source? Are there any sources at all? Who are the stakeholders? Do yourself the courtesy of not sharing inaccurate data.

## Works Cited

- Allen, Elena A, et al. "Data Visualization in the Neurosciences: Overcoming the Curse of Dimensionality." *Neuron*, vol. 74, no. 4, 24 May 2012, pp. 603–608., doi:10.1016/j.neuron.2012.05.001.
- Beaulieu, Sarah. "The Truth About False Accusation." *Sarah Beaulieu*, 1 Dec. 2012, [sarahbeaulieu.me/the-truth-about-false-accusation](http://sarahbeaulieu.me/the-truth-about-false-accusation).
- "Data Visualization Beginner's Guide: a Definition, Examples, and Learning Resources." *Tableau Software*, 2019, [www.tableau.com/learn/articles/data-visualization](http://www.tableau.com/learn/articles/data-visualization). (Module 3)
- Esteban, Chiqui. "A Quick Guide to Spotting Graphics That Lie." *National Geographic*, 19 June 2015, [www.nationalgeographic.com/news/2015/06/150619-data-points-five-ways-to-lie-with-charts/](http://www.nationalgeographic.com/news/2015/06/150619-data-points-five-ways-to-lie-with-charts/). (Module 7)
- Estrada, Fabiola Cristina Rodríguez, and Lloyd Spencer Davis. "Improving Visual Communication of Science Through the Incorporation of Graphic Design Theories and Practices Into Science Communication." *Science Communication*, vol. 37, no. 1, 26 Dec. 2014, pp. 140–148., doi:10.1177/1075547014562914. (Module 6)
- "False Reporting Overview." *National Sexual Violence Resource Center*, 2012, [www.nsvrc.org/publications/false-reporting-overview](http://www.nsvrc.org/publications/false-reporting-overview).
- Gleick, Peter. "Global Temperature Anomalies from 20th Century Average 2007 to 2011 (Degrees C)." *Forbes*, 5 Feb. 2012, [www.forbes.com/sites/petergleick/2012/02/05/global-warming-has-stopped-how-to-fool-people-using-cherry-picked-climate-data/#22d4aef42b5c](http://www.forbes.com/sites/petergleick/2012/02/05/global-warming-has-stopped-how-to-fool-people-using-cherry-picked-climate-data/#22d4aef42b5c).
- Hepworth, Katherine. "Big Data Visualization." *Communication Design Quarterly Review*, vol. 4, no. 4, 27 Dec. 2017, pp. 7–19., doi:10.1145/3071088.3071090.
- "Information Visualization – A Brief Introduction." *The Interaction Design Foundation*, 2019, [www.interaction-design.org/literature/article/information-visualization-a-brief-introduction](http://www.interaction-design.org/literature/article/information-visualization-a-brief-introduction). (Module 3)
- Janssen, M. P. "What Does Your Data Tell You? How Transfusion Chain Data Can Support Managerial Decision-Making." *ISBT Science Series*, vol. 12, no. 1, 21 Nov. 2016, pp. 38–45., doi:10.1111/voxs.12323.
- Kirsch, Noah. "The Inside Story Of Papa John's Toxic Culture." *Forbes*, Forbes Magazine, 19 July 2018, [www.forbes.com/sites/forbesdigitalcovers/2018/07/19/the-inside-story-of-papa-johns-toxic-culture/#3ed71a53019f](http://www.forbes.com/sites/forbesdigitalcovers/2018/07/19/the-inside-story-of-papa-johns-toxic-culture/#3ed71a53019f).
- Lebied, Mona. "Misleading Statistics & Data – News Examples For Misuse of Statistics." *Datapine*, 8 Aug. 2018, [www.datapine.com/blog/misleading-statistics-and-data/](http://www.datapine.com/blog/misleading-statistics-and-data/). (Module 7)
- Macauley Millar, Thomas. "Meet The Predators." *Yesmeansyesblog.wordpress.com*, 16 Nov. 2009, [yesmeansyesblog.wordpress.com/2009/11/12/meet-the-predators/](http://yesmeansyesblog.wordpress.com/2009/11/12/meet-the-predators/).
- Marsh, Samantha. "When Data Visualizations Mislead (And How to Prevent It)." *IDashboards*, [www.idashboards.com/blog/2019/02/27/when-data-visualizations-mislead-and-how-to-prevent-it/](http://www.idashboards.com/blog/2019/02/27/when-data-visualizations-mislead-and-how-to-prevent-it/).
- McArdle, Megan. "Ending the Infographic Plague." *The Atlantic*, Atlantic Media Company, 24 Dec. 2011, [www.theatlantic.com/business/archive/2011/12/ending-the-infographic-plague/250474/](http://www.theatlantic.com/business/archive/2011/12/ending-the-infographic-plague/250474/).
- McCormack, David. "CNN Reports 110% Turnout in Scottish Independence Vote." *Daily Mail Online*, Associated Newspapers, 18 Sept. 2014, [www.dailymail.co.uk/news/article-2761778/Something-doesn-t-add-CNN-Reports-110-turnout-Scottish-independence-vote.html](http://www.dailymail.co.uk/news/article-2761778/Something-doesn-t-add-CNN-Reports-110-turnout-Scottish-independence-vote.html).
- Miller, Paul, and Julian Richards. "The Good, the Bad, and the Downright Misleading: Archaeological Adoption of Computer Visualization." *Computer Applications and Quantitative Methods in Archaeology*, 1995, pp. 19–22.,

proceedings.caaconference.org/files/1994/03\_Miller\_Richards\_CAA\_1994.pdf.

Reuters. “Gun Deaths in Florida.” *LiveScience*, 16 Feb. 2014, [www.livescience.com/45083-misleading-gun-death-chart.html](http://www.livescience.com/45083-misleading-gun-death-chart.html).

Robbins, Naomi. “Misleading Graphs: Figures Not Drawn to Scale.” *Forbes*, Forbes Magazine, 16 Feb. 2012, [www.forbes.com/sites/naomirobbins/2012/02/16/misleading-graphs-figures-not-drawn-to-scale/#49ccb9bd15ef](http://www.forbes.com/sites/naomirobbins/2012/02/16/misleading-graphs-figures-not-drawn-to-scale/#49ccb9bd15ef).

Sloan, Christopher. “Visual Storytelling for Science.” *Science Visualization*, 2019, [sciencevisualization.com/](http://sciencevisualization.com/). (Module 6)

Yang, Chaowei, et al. “Big Data and Cloud Computing: Innovation Opportunities and Challenges.” *International Journal of Digital Earth*, vol. 10, no. 1, 3 Nov. 2016, pp. 13–53., doi:10.1080/17538947.2016.1239771.

Zen, Pola. “Storytelling Secrets For Creating Images That Connect.” *Yotpo*, 27 Dec. 2017, [www.yotpo.com/blog/5-visual-storytelling-secrets-to-improve-your-marketing-images/](http://www.yotpo.com/blog/5-visual-storytelling-secrets-to-improve-your-marketing-images/). (Module 5)

SHARE:

RATE:

Previous

Next

Where did you get that image? The greatest sin in visual storytelling

I Am My Content: There’s Magic in That Relationship

ABOUT THE AUTHOR

RELATED POSTS



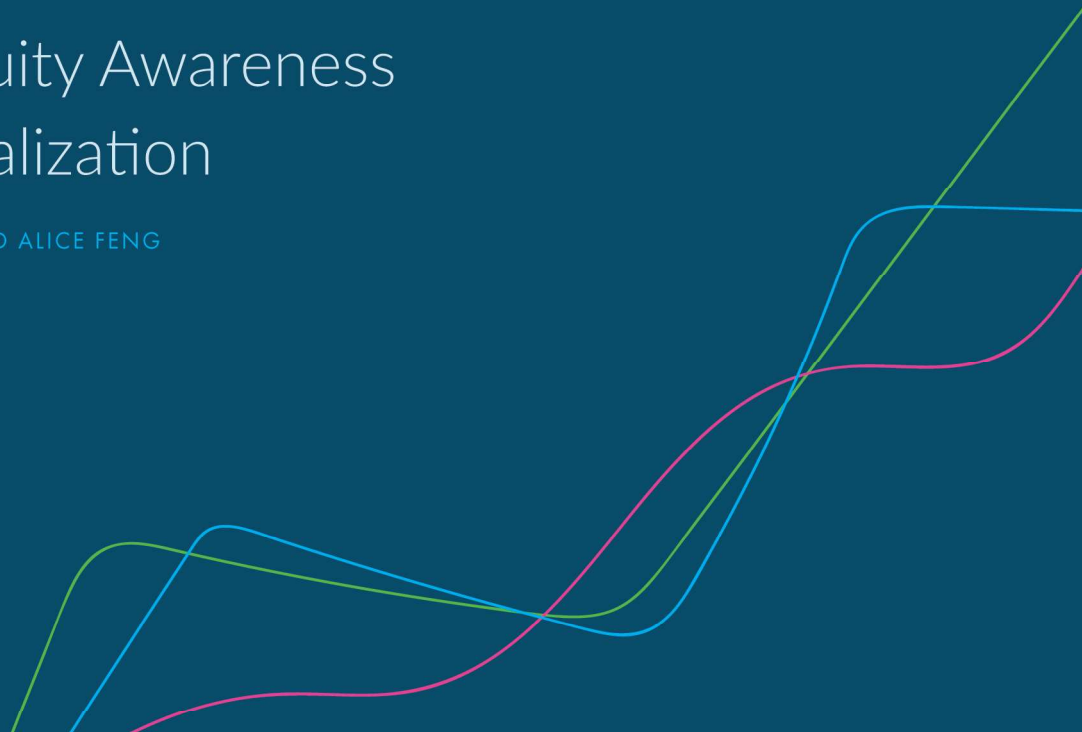


# DO NO HARM GUIDE

## Applying Equity Awareness in Data Visualization

JONATHAN SCHWABISH AND ALICE FENG

JUNE 9, 2021



## TABLE OF CONTENTS

---

01 INTRODUCTION	3	12 BUILDING DIVERSE AND INCLUSIVE DATA COMMUNICATION TEAMS AND ORGANIZATIONS	26
02 RACE AS A CONSTRUCTION	4	13 THE FEEDBACK LOOP	27
03 DEMONSTRATING EMPATHY	5	14 THE GATEKEEPERS	28
04 ENGAGING OR REFLECTING LIVED EXPERIENCES	9	Editors and Peer Reviewers	28
Methods	10	Government Agencies, Regulations, and Directives	29
05 THE DATA BEHIND THE VIZ	11	Funders	30
Data Confidentiality and Security	12	15 CASE STUDY FROM THE URBAN INSTITUTE	30
06 USING LANGUAGE WITH A RACIAL EQUITY AWARENESS	13	16 LOOKING FORWARD	33
07 ORDERING DATA PURPOSEFULLY	15	17 ACKNOWLEDGMENTS	34
08 CONSIDERING THE MISSING GROUPS	17	18 REFERENCES	35
Lumpers and Splitters	17	19 INTERVIEW GUIDE: DO NO HARM GUIDE AGENDA	39
Male, Female, In Between, and Beyond	19	20 OTHER DATA AND DATA VISUALIZATION RACIAL EQUITY RESOURCES	40
Othering the Other	20	21 DIVERSITY, EQUITY, AND INCLUSION IN DATA VISUALIZATION: GENERAL RECOMMENDATIONS	41
Choosing to Not Include All Groups	21	22 THE RACIAL EQUITY IN DATA VISUALIZATION CHECKLIST	42
09 USING COLORS WITH A RACIAL EQUITY AWARENESS	22		
10 USING ICONS AND SHAPES WITH A RACIAL EQUITY AWARENESS	23		
11 EMBRACING CONTEXT AND COMPLEXITY	25		

---

## 01 INTRODUCTION

People who work with data every day seek to make discoveries, elicit insights, and offer solutions. To do so in an equitable way, data analysts should think intentionally about how we can learn from and speak to audiences that reflect the rich diversity of the people and communities of focus. Failing to recognize the diversity in these groups will further exacerbate—and potentially contribute to—the inequities that have shaped the world. Today, in the era of big data, data visualization, machine learning, and artificial intelligence, we need to be more purposeful about where data are coming from and how research and the communication of that research can affect people, their communities, and the policies that touch their lives.

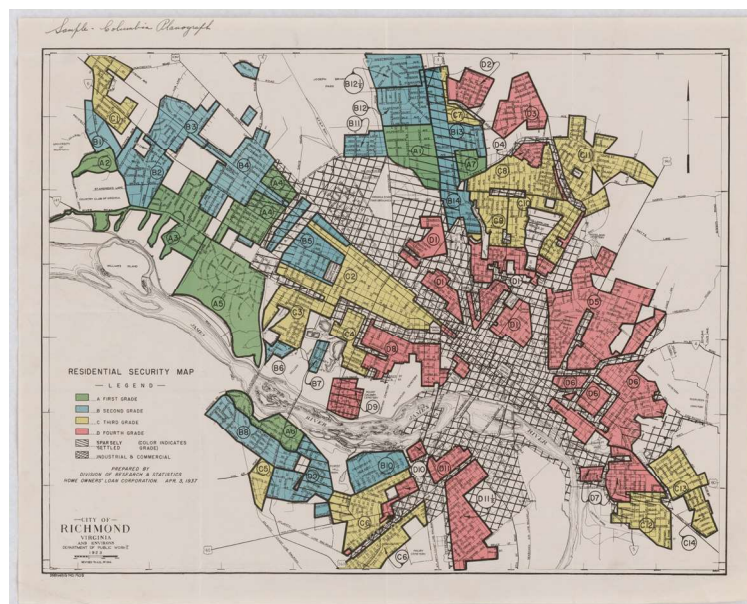
Personal and organizational efforts to foster diversity, equity, and inclusion (DEI) need to extend not only to our internal processes of hiring, promotion, and professional development but also to our external communication efforts to represent the people and communities we focus on. To that end, we need to consider how to apply a DEI lens not just to the words, colors, icons, and other elements in our writing, graphs, charts, and diagrams,

but also to the process of crafting these communication products.

This guide and the associated checklists and toolkits focus on the often hidden or subtle ways that data analysts and communicators fail to incorporate equitable awareness in the data they use and the products they create. Some data and data visualizations, however, are overtly racist and discriminatory. For example, the map shown in figure 1 was produced in 1937 by the Home Owners' Loan Corporation (HOLC), a federal agency tasked to appraise home values and neighborhoods across the United States. Richard Rothstein described the HOLC and these maps in his book, *The Color of Law*: “The HOLC created color-coded maps of every metropolitan area in the nation, with the safest neighborhoods colored green and the riskiest colored red. A neighborhood earned a red color if African Americans lived in it, even if it was a solid middle-class neighborhood of single-family homes” (Rothstein 2017). The bulk of this guide does not focus on data and visuals that are overtly racist, sexist, or discriminatory, but as data practitioners, we still need to

FIGURE 01

This redlining map of Richmond, Virginia, from the Home Owners' Loan Corporation shows how data and data visualization can be used to further systematic discrimination.



Source: National Archives.

be aware of how such work often creates irreparable and long-term harm. Systematic discrimination is and can be generated by how we use and misuse our data.

In this guide, we explore ways to help data scientists, researchers, and data communicators take a more purposeful DEI approach to their work. To do so, we conducted more than a dozen interviews with nearly 20 people about their experiences and approaches to being more inclusive with their data exploration, analysis, and communication. Our interviewees included data journalists in major media outlets, researchers at universities and colleges, and people working in the public and private sectors. Our long (and growing) list of people we would have liked to speak with for this project indicates how the work of creating diverse and inclusive products in the data and data visualization areas will continue to grow. We also take a broad view of the “data community” to include anyone working with and communicating data: researchers, scholars, data analysts and scientists, journalists, web developers, data visualization specialists, and more.

Our recommendations are not one-size-fits-all and should be adapted to the particular needs of a project and organization. A large research project with a team of researchers that spans multiple months or years might embrace some of these outreach efforts differently than a single data visualization intended for a 600-word blog post or a standalone dashboard. Through honest conversation and evaluation of different projects, data project creators can be better prepared to incorporate the strategies described here in future work.

The motivation for this entire project can be summarized in one word: **empathy**. Applying a DEI lens to how we analyze, visualize, and communicate data requires empathizing with both the communities whose data we are visualizing as well as the readers and target audiences for our work. This means considering how the lived experiences and perspectives of our study populations and readers affect how they will receive and perceive the information. This quote from journalist Kim Bui succinctly captures the concept: “If I were one of the data points on this visualization, would I feel offended?” As we consider the use of words, colors, icons, and more in our data visualizations, asking whether we would be offended makes for a good checkpoint.

---

### **If I were one of the data points on this visualization, would I feel offended?**

– KIM BUI

---

This guide offers a set of guidelines rather than ironclad rules for presenting data through a DEI lens. Our goal is not to tell data analysts what to do or not to do but to instead ask them to be thoughtful in the ways they work with and communicate their data and to be aware of the decisions they have made and why they have made them. As we will show, the choices we make about the colors, shapes, words, and representations in our data analysis and visualizations can affect how people perceive the final results, how change might be implemented, and how that change will impact different people and communities.

## **02 RACE AS A CONSTRUCTION**

---

In this guide, we focus primarily on race and ethnicity, but similar issues of inclusivity around gender, sexuality, class, ability, and other characteristics (as well as their intersectionality) should always be at the forefront for every data analyst and communicator. The principles outlined here are also applicable to addressing these other aspects of identity.

With respect to race specifically, it is important to remember that race is an artificial human construct. Racial and ethnic categories such as “Asian,” “Black,” “Hispanic,” “white,” and others are classifications and labels society

places on people, but they are not rooted in biology or genetics. “Race,” writes scholar Ibram X. Kendi (2019) in *How to Be an Antiracist*, is “a power construct of collected or merged difference that lives socially” (35). Later in the book, Kendi writes:

...race is a mirage, which doesn’t lessen its force ... What people see in themselves and others has meaning and manifests itself in ideas and actions and policies, even if what they are seeing is an illusion. (37)

Thus, as communicators consider how to analyze and visualize issues of race, ethnicity, and other characteristics, they should be cognizant of how these categorizations are used and perceived in different fields and areas of study. In the United States, as well as many other countries, the social construct of race has been, and continues to be, used to create and maintain systems that establish the power and privilege of certain groups (i.e., white people) while discriminating against and oppressing other groups (i.e., Black, Hispanic, Native American, Asian, and other people). Such racism does not just operate at the individual and interpersonal levels but is baked into institutions, laws, and norms, and thus it manifests in the form of institutional and structural racism that perpetuates these benefits and harms across generations. This has resulted in lasting impacts among people and communities of color, such as the lower levels of wealth held among Black and Hispanic families (McKernan et al. 2017) and higher rates of chronic health conditions among people of color (Carratala and Maxwell 2020). As Isabel Wilkerson (2020) writes in her book *Caste*, “The

human impulse to create hierarchies runs across societies and cultures, predates the idea of race, and thus is farther reaching, deeper, and older than raw racism and the comparatively new division of humans by skin color” (67).

When thinking about how to analyze and communicate data with a racially equitable lens, we should ask ourselves whether the issue we are analyzing and visualizing is worse or exacerbated for people of color and what factors contribute to or compound racial inequities for this issue. These issues have taken even greater urgency and focus with the announcement from the Biden administration that the federal government “should pursue a comprehensive approach to advancing equity for all, including people of color and others who have been historically underserved, marginalized, and adversely affected by persistent poverty and inequality” (White House 2021). It is important to be informed and aware of the policies, institutions, and actors that have shaped such inequities and disparities and to acknowledge this history and context in our communication products.

## 03 DEMONSTRATING EMPATHY

### **You can't feel empathy for a point on a map, but you can feel empathy for a mom who's crying.**

– KIM BUI

One of the big challenges in visualizing data, and of quantitative research in general, is the ability to help readers connect with the content. Merriam-Webster defines empathy as “the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner.” Connecting directly with people and communities and trying to better understand their lived experiences can help content producers create visualizations and tell stories that better reflect those people's true experiences.

This is a key dimension for people working with data to explore to help put their work into the hands of policymakers, stakeholders, and community members who can use it to effect change. Inclusive and thoughtful data visualizations that respectfully reflect the people we are studying can help us build trust with those communities. As Kim Bui (2019) wrote, “approaching stories—and people—with more empathy creates better relationships with marginalized communities, builds trust and increases diverse coverage.”

We think of empathy as it applies to communicating data across six main themes:

#### **1. Put people first.**

First and foremost, we need to remember and communicate that the data shown reflect the lives and experiences of real people. Data communicators must help readers better understand and recognize the people behind the data. As Jacob Harris (2015) wrote, “If your data is about people, make it extremely clear who they are or were.”



## 2. Use personal connections to help readers and users better connect with the material.

As Kim Bui told us, “You can’t feel empathy for a point on a map, but you can feel empathy for a mom who’s crying.” Pairing data-driven charts with personal stories centered on individual experiences can help readers understand and identify with the people represented in the research and data visualizations. Techniques that can be used in tandem with data visualizations to help lift up personal stories include photography, illustrations, pull quotes, and oral histories. Helping readers understand the “near and far” (Harris 2015)—a wide lens of overall metrics or data paired with individual- or group-level data—may help readers connect with content (Klein, n.d.).

Making charts relatable can also be important for audience engagement and impact. In their study of what factors influenced attitudes and perceptions about data visualizations in rural populations, Evan Peck and coauthors found that visualizing topics that the audience had personal connections to trumped design factors when rating which charts they liked the most. They write, “Regardless of style, clarity, or ease-of-understanding, our interviews served as a reminder that data can be intimate and personal, and that those ties may supersede many other dimensions of design” (Peck, Ayuso, and El-Etr 2019). Similarly, allowing users to find themselves in the visualization by, for example, being able to search for their city or county in a map, can also strengthen audience engagement with the data. There is an important balance here, of course, between being able to search and filter for a specific geographic area or subpopulation and data privacy and confidentiality, which we address in section 5.

## 3. Use a mix of quantitative and qualitative approaches to tell a story.

Most charts and graphs are built on top of spreadsheets or databases of quantitative data. However, focusing on numbers alone without any context can overlook important aspects of a story including the “why” and the “how.” We discuss other research methods that analysts can use in section 4.

## 4. Create a platform for engagement.

This can take the form of interactivity in which users are able to manipulate buttons, sliders, tooltips, and other elements to make selections, filter the dataset, or create customized views of a chart. Such engagement can be leveraged as a way to allow users to find

themselves in the data or discover the stories that most interest them. Another form of engagement is offering audiences a means of providing feedback about a data tool or visualization. Feedback forms offer users a way to share their own experiences with the topic being visualized. We discuss this further in section 12.

## 5. Consider how your framing of an issue can create a biased emotional response.

In our discussion with Kim Bui, she pointed to coverage and discussion about racial justice protests in the spring and summer of 2020. Much of that coverage, she argued, tended to focus on the number of riots that occurred or how much property damage took place, which portrayed the protests and protesters in a negative light while missing the cause of those protests: “people trying to get more justice for Black and brown lives that have been taken from us.” Similarly, in her article “When the Designer Shows Up in the Design,” Lena Groeger discusses two different ways to visualize the impact of crime on local communities: maps showing the locations where crimes occurred versus maps showing the percentage of residents in a neighborhood who were in prisons. The former focus on victims of crimes while the latter “show the urban costs of incarceration and suggest how those dollars might be better spent on investing in communities” (Groeger 2017). What data we choose to focus on and what we choose to ignore can bias our audiences’ perceptions of the issues we are communicating about.

## 6. Recognize the needs of your audience.

Taking an empathetic view of the readers’ needs as they read or perceive information is also an important step to better data communication (Schwabish 2021). This kind of empathy is often couched in terms of producing specific graphs that meet the needs and expertise of our readers. For example, nearly 28 million people in the United States alone have challenges accessing content on the web, ranging from issues of sight to physical or intellectual impairments (Brault 2012). Making sure our visualizations are accessible is important so people with disabilities can view and use online content (see also Lundgard, Lee, and Satyanarayan 2019). Similarly, overly technical or jargon-laden language can make reports and articles inaccessible to broader audiences and less likely to be cited by other researchers in the field (Martínez and Mammola 2021). Another consideration is translating materials into languages most used by your audiences.



Being empathetic to the people and communities of focus does not imply sacrificing the data and methods in responsible, in-depth, sophisticated research. In fact, the opposite is true: research and analysis must be responsible and as objective as possible to create effective policy solutions and recommendations. We are not arguing that empathy replaces high-quality research or data work; rather, such research and empathy for people and communities can be complementary. Effective research necessarily means understanding someone else's point of view nonjudgmentally and recording that perspective as accurately and truthfully as possible. Engaging in empathy allows researchers and data analysts to better understand the source and full context of the data they are working with. (Given the considerations around bias in data and in human decisionmaking generally, we hesitate to use the term "objective" as it applies to research, but that is essentially what we are trying to imply here—that empathy and objectivity in research are not opposites but complements for generating more effective research and policy solutions.)

In thinking about embedding empathy into the design of data visualizations, it is worth considering whether certain chart types better lend themselves to connecting with the human element of the data being portrayed. Commonly used charts and graphs like bar charts, line charts, and pie charts, while informative, can abstract the content by collapsing all the people represented into one shape.

**These are not just data points, and so when you add them as dots on a map, or lines on a chart, I think you have to remember that these are true people.**

– TIM MEKO

Unit charts and dot plots, which use multiple repeated shapes to depict the data, might offer more opportunity to connect with the subject by reminding readers of the number of people represented, particularly if each dot represents one person.

Taking this thinking one step further, the use of icons instead of abstract shapes such as circles and rectangles may also improve the ability of readers to empathize by reinforcing that they are looking at people and not just numbers or statistics. As Tim Meko from the *Washington Post* told us, "These are not just data points, and so when you add them as dots on a map, or lines on a chart, I think you have to remember that these are true people." Graphics that specifically represent people—the anthropomorphizing of data graphics, or, as Jeremy Boy and colleagues refer to it, anthropographics (Boy et al. 2014)—is sometimes seen as a way to evoke empathy (though Boy and his coauthors do not find this to be the case; see also Groeger 2014). An example of a stacked bar chart compared with a unit chart that uses icons is shown in figure 2.

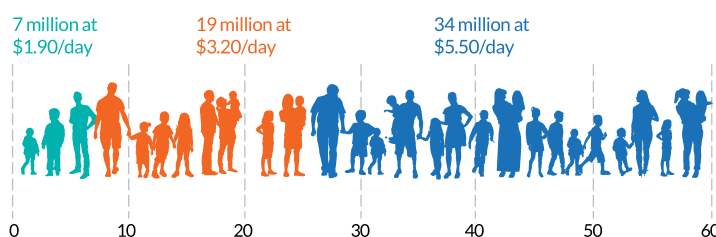
FIGURE 02

A comparison of a standard stacked bar chart and a unit chart with icons.

### Millions of people in poverty



### Millions of people in poverty



Source: Recreated based on Schwabish (2021).

In a paper examining the design space of anthropographics, Luis Morais and authors articulate seven design dimensions that may affect the ability of a data visualization to elicit empathy or compassion from the audience (Morais 2020). These include *granularity* (the number of people represented by one shape or graphical mark), *specificity* (how completely each item in a chart can be distinguished from another), *coverage* (the number of people visualized), and *authenticity* (whether the chart shows any additional attributes that are not derived from the source dataset). In addition to what and how much information is shown, how that information is represented is also worth considering. This includes the *realism* of the visualization, with abstract shapes such as bars and lines being the least realistic and photographs being the most; *physicality*, meaning the use of physical versus virtual objects; and *situatedness*, or how physically close or far a visualization is from the people or setting it describes. Although current research does not yet offer guidance on the effectiveness of these dimensions, they may be aspects worth considering when designing your data visualizations.

Using human icons to visualize data does not necessarily guarantee empathy, however. The chart in figure 3 has been ridiculed as an example of how not to use icons. The icons are scaled inappropriately by both height and width, the y-axis does not start at zero, and the icons use stereotypical pink colors. The scaling distorts the representation of the data, and the resulting chart has been perceived as offensive to many women, particularly those in the six countries portrayed in the chart. Sabah Ibrahim, an educator and environmental campaigner, shared this tweet in August 2020: “As an Indian woman, I can confirm that too much of my time is spent hiding behind a rock praying the terrifying gang of international giant ladies and their Latvian general don't find me.” While the tweet is somewhat tongue-in-cheek, it hearkens to the initial motivation in this paper: “If I were one of the data points, would I feel offended?”

The issue of empathy and how data communicators should take an empathetic view to the data they are communicating and the people they are communicating

about underlies the concept of conducting research and making data visualizations that uphold DEI. In the case of racial equity, we can take that concept of empathy a step further and recognize the histories of racism, injustice, abuse, and discrimination that serve as the background for many of the current issues we analyze, visualize, and debate.

FIGURE 03

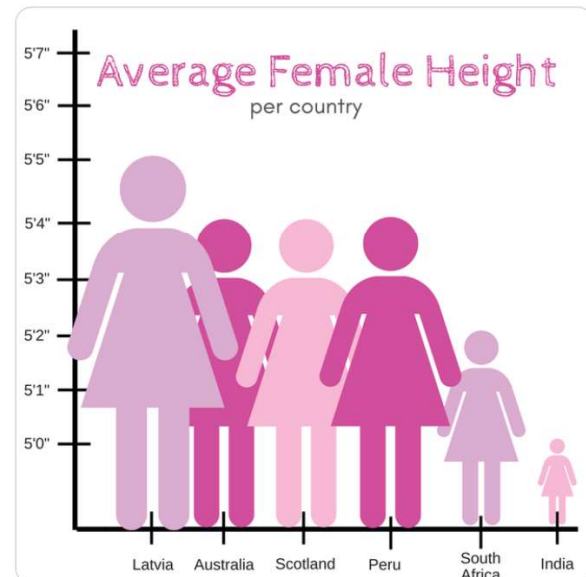
A tweet from Sarah Ibrahim demonstrates how people can see themselves in the data.



Sabah Ibrahim  
@reina\_sabah

...

As an Indian woman, I can confirm that too much of my time is spent hiding behind a rock praying the terrifying gang of international giant ladies and their Latvian general don't find me



6:58 PM · Aug 6, 2020 · Twitter Web App

Source: Sarah Ibrahim (@reina\_sabah), “As an Indian woman, I can confirm that too much of my time is spent hiding behind a rock praying the terrifying gang of international giant ladies and their Latvian general don't find me,” Twitter, August 6, 2020, 6:58 p.m., [https://twitter.com/reina\\_sabah/status/1291509085855260672?lang=en](https://twitter.com/reina_sabah/status/1291509085855260672?lang=en).

## 04 ENGAGING OR REFLECTING LIVED EXPERIENCES

---

### **Don't assume, ask people.**

– KIMBERLYN LEARY

---

The need for data communicators to understand and form relationships with the people and communities that are the focus of their research emerged as a common theme in all of our discussions. Research and analysis should be done with communities, not on behalf of them, except in the rare cases where doing so is not feasible or appropriate. Such communities should derive value from the research and data and have some level of input into the design of the project itself and the final deliverables. This kind of buy-in—from members of the community, policymakers, and other stakeholders—can help the research be more impactful, relevant, and embraced by a wider audience.

We do not, however, intend to suggest that research objectivity and independence be sacrificed to engage with communities. We also recognize that interacting with people and communities does not work for every research project or data visualization: national, aggregated data from the federal government, for example, might be less appropriate for these approaches than projects involving first-hand data collection. But for projects with opportunity for such engagement, there is likely a middle ground where community members can weigh in and feel heard and recognized.

In light of these considerations, researchers and analysts should reach out to and engage communities from the start of the project. This is an important step for making sure a DEI framework and empathy are baked into the project and kept in mind from the beginning rather than “retrofitted” at the end. Reaching out to local organizations and nonprofits, service providers, social justice groups, and advocacy groups, as well as to your personal network for connections or searching the internet for people talking about the topic you are interested in, can help identify community partners. Some universities and newsrooms have engagement teams or offices that specialize in working with local organizations; such teams can also be a source for forming connections with communities.

A critical aspect to keep in mind is understanding a community's culture, history, and values. Different communities have different preferences for how they wish to be approached, and it is important to respect those preferences. For example, members of the University of California Los Angeles (UCLA) Center for Health Policy Research told us that Native Hawaiian Pacific Islander communities prefer to meet face to face and have the researchers go to them rather than asking community members to come to the researchers. Relationships, they argued, need to be built before you can ask for a favor (see also Costanza-Chock 2018).

After establishing such relationships, researchers and analysts should seek to understand the needs of the community and how the work they are doing would benefit them, engaging them as partners early on in the process. Kimberlyn Leary, senior vice president at the Urban Institute, recommended that researchers “mak[e] sure that you put in the time to understand not just the disparities, but engage with people to find out and offer some solutions... don't assume, ask people.” Kim Bui echoed that sentiment, saying it is important to tell stories with a community rather than on behalf of them and to seek out what they would want to learn and what would be useful for them along with their concerns. Asking community members for their perspectives can also help researchers better understand the context surrounding the issue they wish to study as well as how the community perceives that issue. Similarly, community members can help interpret the data, but which community members (or groups) researchers should speak to and how many people they should speak with are important considerations, especially when research funds are limited or there is significant variation within the community. Those issues underscore the importance of engaging with communities at the beginning of the research project.

While engaging with communities, researchers and analysts should be clear and transparent up front on what their intentions are and what they will be doing with the data they gather. Where possible and appropriate, the project's goals should align with the community's goals, and benefits should ideally be tangible early on in the

project cycle rather than occurring at the end. Research questions should reflect the community's values and account for cultural and historical context. Community members' input should be sought to ensure that the design of the research upholds these goals (but again, the data used and methods applied should be rooted in the science objectively and honestly). Above all, researchers should not exploit people and their experiences by taking data from the community and not giving back. Community members should not only benefit from the results of the work but also have a vested interest in the research, data, and final products. However, because not all community members will agree on certain findings or policy recommendations, the ultimate responsibility for determining the final conclusions still lies in the hands of the researchers.

As the project nears the end, results, data, and final products should be shared with the community first before being released to the public. Community members should be given a chance to review and provide feedback before publication. Questions to ask about communications products include the following: Are we using language consistent with how community members refer to themselves and others? What in our visuals is inconsistent with a DEI framework? And how does the community prefer to have the research results presented? The feedback received should then be critically reviewed by the research team and incorporated into final visualizations and products as appropriate, and community members should be credited for their contributions.

Finally, once built, these relationships should be sustained. This is important for fostering and maintaining trust so the community does not feel like it is being used or taken advantage of and to ensure that the community is able to act on the recommendations that result from the research. Such relationships can also spur collaborations and partnerships for future projects.

## **METHODS**

Quantitative researchers and analysts especially should consider how best to incorporate qualitative methods when conducting research that engages lived experiences. Long-form surveys, interviews, and focus groups can provide an important opportunity for community members to share their experiences

and lift up their voices. Such questions can also help answer "why" questions and surface themes in the responses. Evan Peck from Bucknell University, for example, interviewed 42 people in rural Pennsylvania to better understand their relationships to data and data visualizations (Peck et al. 2019). Peck and coauthors found that people's reactions to data visualizations were often driven or framed by their personal experience (e.g., where they live or work, their level of educational attainment, or their political identity), suggesting that our existing knowledge of how data visualizations facilitate understanding is much broader than whether and how people can discern quantities from the graph (Cleveland and McGill 1984; Heer, Bostock, and Ogievetsky 2010). Their conversations with study participants enabled them to better understand the detailed explanations behind their reasoning about how they read and understand graphs that more standard quantitative methods may not have permitted. In our interview with Peck, he elaborated further and noted that he suspects his team had higher participation in their interviews because participants could see how the data were being collected (rather than being collected on an impersonal form), so they had "some sort of understanding of how the data is collected."

More generally, our interviews suggest that embracing a "reciprocal research" strategy—in which research participants see concrete, actionable benefits—makes research stronger or more accessible and its recommendations more likely to succeed (Myers and MacDonald 2021). With data visualization specifically, some have argued that information visualization should empower users to allow them to "question visual representations, utilize them to tell their own story, and shift from awareness to action" (Dörk et al. 2013, 7). Dr. Sasha Costanza-Chock at the Massachusetts Institute of Technology has helped popularize the Design Justice model that also recognizes that community participation is not singular nor necessarily consistent during the project. "The point" Costanza-Chock writes, "is to encourage a more complex understanding of participation and to emphasize that very few design processes are characterized by community control throughout" (Costanza-Chock 2018).

Again, these approaches will not work for every research project and every data visualization. Many projects will not interact with people directly or will address

aggregate concepts affecting numerous groups, and many visualizations will be geared toward a specific purpose for a specific audience. But research and data communities should at least consider these issues as they progress through the project.

We conclude this section with a final caveat about these kinds of qualitative methods and research: we are not suggesting that pursuing qualitative research is as simple as conducting a few interviews. Entire fields are devoted to qualitative methods, which require study, training, and practice to master. In our experience, we have found that

many primarily qualitative researchers have at least some quantitative training: they know how to clean quantitative data, estimate summary statistics, and run regressions. By comparison, many quantitative researchers have little concept of what it means to conduct qualitative research—the methods, skills, and practices to effectively collect and understand such data using these kinds of techniques. We hope that this guide can help move qualitative and quantitative research fields closer to one another.

## 05 THE DATA BEHIND THE VIZ

---

Before we jump into discussing ways to apply racial equity awareness to charts and maps, we should first step back and critically examine the data underlying them. Just as trying to analyze flawed data generates flawed results, trying to visualize flawed data will also generate misleading and deceptive charts. When thinking about DEI issues specifically, we should check our data for ways they may be biased or the product of racist or oppressive data collection systems. As Evan Peck noted, data visualizations amplify data by making them more accessible. Thus, if the data underlying the visualization are biased, the chart will amplify that bias and the harm that bias generates.

A full discussion of how to identify and handle biased data is beyond the scope of this guide, but the questions below offer a good starting point for what data analysts and communicators should be asking when they receive a dataset (for more ideas, see our list of additional resources provided at the end of this guide):

- How were these data generated?
- Are these data demographically representative?
- Who is included and who is excluded from these data?
- Whose voices, lives, and experiences are missing?
- How much can this data be disaggregated by race, gender, ethnicity, etc.?
- Why were these data collected?
- Who stands to benefit from these data?
- Who might be harmed by the collection or publication of these data?

Data, particularly data collected about people, are not neutral or objective: they reflect a human act that was done by someone for a specific purpose and intent (Correll 2019; Lupi 2017). Data visualizers and communicators must first thoroughly understand the data they are using, how those data were collected, why they were collected, and who is and is not reflected in those data before they begin creating charts and graphics. We strongly emphasize this point because applying DEI thinking to visualizations alone will not fix the problems of data or analyses that are biased or racist. Researchers and analysts should also consider the root causes of disparities that are revealed in the data and frame their work and interpretations in light of that knowledge.

Taking a racially equitable perspective to data should be baked into a project from the beginning and continue throughout the process. This entails critically examining how the study is designed and how the data are collected as well as how the analysis of that data is being conducted and the results communicated. If the data or analysis are inherently racist or biased, applying the principles in this guide to the way you visualize these data will not magically make your visualization “anti-racist” (Kendi 2019). It can also be valuable to think about the impact of racism and other forms of oppression on the data you are visualizing even if your charts do not explicitly show racial or ethnic breakdowns. DEI thinking cannot be a Band-Aid slapped on at the end to fix a project that is inherently racist or perpetuates oppression.



## DATA CONFIDENTIALITY AND SECURITY

Two areas of working with data do not always lend themselves to full and transparent data communication: uncertainty, and privacy and security. Uncertainty in data adds an important challenge to all data analysis and is one that affects not only how data analysis is performed but also when and how uncertainty is presented in data communication (Hullman 2019). Certain audiences are familiar with measures of uncertainty and dispersion, such as confidence intervals, standard errors, percentiles, and variances; other audiences are less fluent in these concepts. Thus, recognizing your audience and providing appropriate labels and annotations to help them better understand these concepts can be central to conveying issues of uncertainty and distribution. Certain conclusions from uncertain evidence can be just as bad as incorrect conclusions from certain evidence. For this reason, good analysis should try to include estimates and measures of uncertainty. Failure to properly communicate uncertainty is a problem with many data projects and can become an even bigger problem when using disaggregated data, especially owing to small sample sizes. Data communicators should explore new and different ways

to communicate uncertainty to their audiences, such as different plot types, better annotation, and even different tools to help users interact with data (Hullman 2019; Nguyen et al. 2020).

With respect to data confidentiality and security, we recognize that many recommendations in this report—such as talking to members of the community or enabling users to see themselves in the data—do not necessarily align with data privacy and security considerations, especially when data are disaggregated. Publishing detailed, disaggregated data can help policymakers and decisionmakers better target policy solutions, but we must carefully consider the unintended harms such detailed data can cause to certain people and groups. A recent report from the Urban Institute notes that “people of color with low incomes are more susceptible to privacy attacks because of their higher reliance on smartphones for internet access and how much personal information they give up for free cell phone app services. This information collection makes them more easily identifiable, especially if they are outliers in small geographies” (Bowen, Williams, and Narayanan 2021; see also Madden 2017 and Bowen 2020).

## 06 USING LANGUAGE WITH A RACIAL EQUITY AWARENESS

### **Titles do such important work to frame the visual part.**

– CATHERINE D'IGNAZIO

Titles, text, and labels are among the first things readers scan when encountering a chart (Borkin et al. 2015), thus presenting an important opportunity to apply racial equity awareness thinking. When possible, forces of oppression such as racism, how those forces are operating, and their historical context should be named directly in the chart through titles, annotations, labels, and notes rather than buried in surrounding text.

The importance of using concise, active titles in data visualizations should not be underestimated. “Titles,” Catherine D'Ignazio mentioned in our conversation, “do such important work to frame the visual part.” Thus, they present an important opportunity to address the

role racism and other forms of oppression play in the topic the chart is visualizing. In their book *Data Feminism* (2020), D'Ignazio and coauthor Lauren Klein offer an example of a bar chart showing the rate of mental health diagnosis of incarcerated people by race (figure 4). The initial title for the chart is “Mental Health in Jail: Rate of mental health diagnosis of inmates.” Although this title seems neutral and objective, it ignores the role racism and discrimination play in how likely incarcerated people are to receive a mental illness diagnosis; it also uses the term “inmate,” which some have argued is dehumanizing and references people by their crimes and punishments (Bartley 2021). The authors then show a different title for the chart: “Racism in Jail: People of color less likely to get mental health diagnosis.” They argue that this alternative title more accurately reflects the main findings of the research (which focused on racial disparities in the jail system), names the forces of oppression at work (racism in prison), and references people, not inmates.



Another alternative takes the process a step further. In figure 5, the subtitle is changed to “White people get more mental health diagnoses,” which focuses not on what people of color lack, but as D'Ignazio and Klein (2020) describe, on the “unfair advantages that are given to the dominant group.” “Placing numbers in context,” they write, “and naming racism or sexism when it is present in those numbers should be a requirement—not only for feminist data communication, but for data communication full stop.”

In addition to the way language can be used to frame the larger picture and context of a chart, it is important to be aware of how language is used in the more detailed parts of a chart, such as labels. Labels should use people-first language, such as “people with disabilities” rather than “disabled people,” or “people in prison” instead of “inmate,” (Bartley 2021) and they should refer to people, not to their skin color (for example, “Black people” rather than “Blacks”).

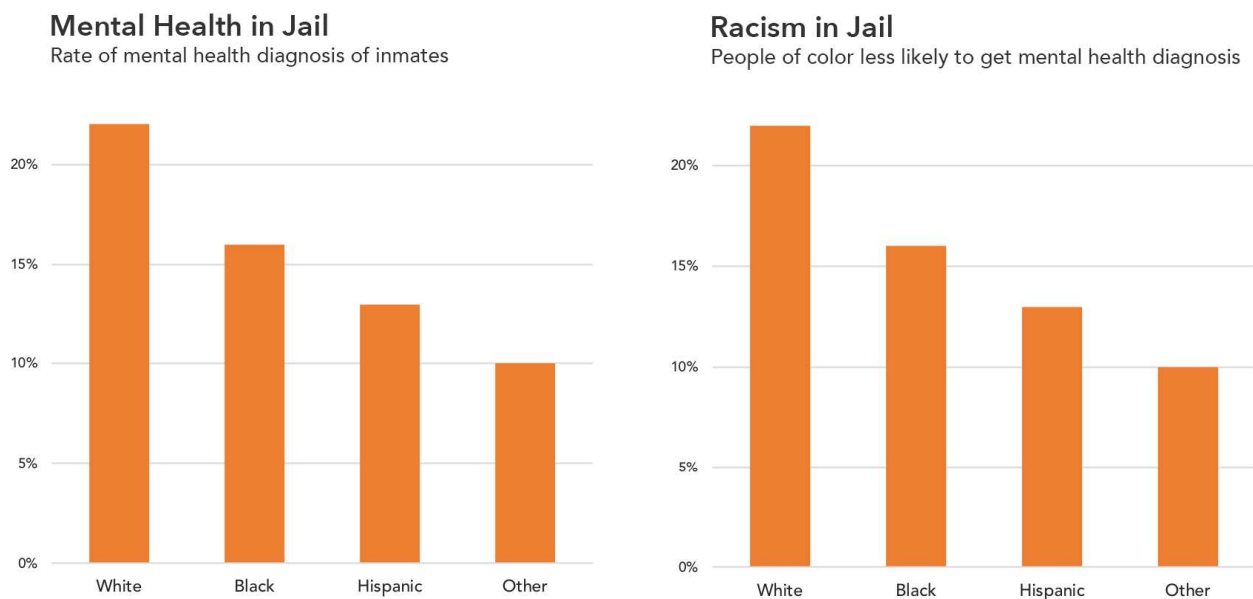
As an example, in a June 2020 project we examined, a series of bivariate choropleth maps showed the relationship between race and poverty. In the original visualization, the labels along the legend were “More Poverty” and “More Black” (figure 6). That language is

not inclusive of different groups: poverty refers to an experience, not a static description; and “More Black” references skin color, not people. A more inclusive way to label the legend might be “Larger proportion of people experiencing poverty” and “Larger Black population” (the author of the visualization later changed “More Black” to “Larger Black Population,” and we are grateful the author provided us with permission to include this example here).

Another challenge when it comes to language is how the terms and phrases used to describe people and communities are constantly changing. In writing about terminology concerning people with disabilities, Nicholas Steenhout (2020) writes, “Disability language is never straightforward. It's always nuanced. It always evolves.” That sentiment can extend to many underrepresented groups; thus, data communicators should monitor the current lexicon and reflect the experiences of the people we study and the people we communicate with. For example, in an August 2020 report, the Pew Research Center (Noe-Bustamante, Mora, and Lopez 2020) found that in their sample of more than 3,000 people who self-identified as Hispanic or Latino, only 23 percent had heard of the term “Latinx” (a gender-neutral alternative that has emerged in the past few years), and

FIGURE 04

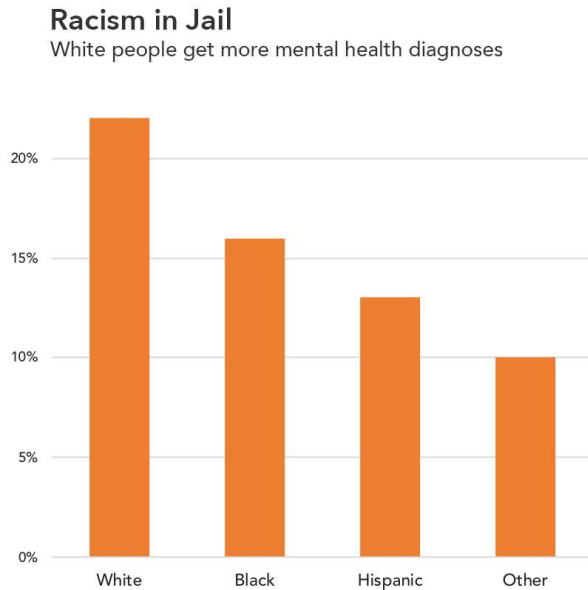
Example of moving from a generic descriptive title to one that reflects the main findings of the research.



Source: D'Ignazio and Klein (2020); Data from Fatos Kaba et al., “Disparities in Mental Health Referral and Diagnosis in the New York City Jail Mental Health Service,” *American Journal of Public Health* 105, no. 9 (2015): 1911–16.

FIGURE 05

Example of a chart title that focuses on the dominant group.



Source: D'Ignazio and Klein (2020); Data from Fatos Kaba et al., "Disparities in Mental Health Referral and Diagnosis in the New York City Jail Mental Health Service," *American Journal of Public Health* 105, no. 9 (2015): 1911–16.

only 3 percent use the term. Of those who have heard the term "Latinx," 65 percent say it should not be used to describe the Hispanic or Latino population.

There may also be generational differences in preferred terms. The same Pew survey found that younger people were more likely to have heard of the term "Latinx" than older generations. G. Cristina Mora captured this well in our discussion by asking whether the data communicator wants to "land future oriented" or "land now." In that sense, she argued, using the term "Latina/o/x" might be "the safest bet or the most inclusive bet" (although there is also a growing movement to use the term "Latine" because it can be easier to pronounce and "-e" can be used with other Spanish words more easily than "-x"; Blas 2019; Del Real 2020.) Given that younger people who identified as Hispanic or Latino might also be more politically progressive, this might also reflect the broader intersectionality of politics, age, and race or ethnicity.

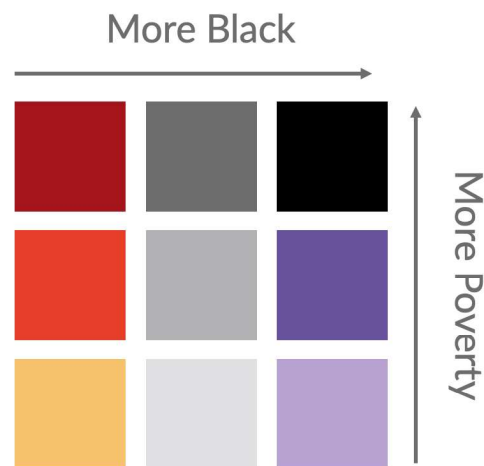
Whichever term the data communicator chooses, one theme that was consistently raised in our discussions was the need for data communicators to talk with their study population and readers to understand and identify

their preferred terminology. Doing so also offers the community whose data were used an opportunity to gain more ownership of the data, research, and policy recommendations.

Finally, we also recognize that surveys or datasets that use outdated or less preferred terminology can present a tricky situation in terms of language. Data communicators may be understandably reluctant to change the word or phrase that was used in the original survey in their final report, graph, or dashboard. If, for example, survey respondents choose the "Hispanic or Latino" option in the survey, is it appropriate for the final product to use a different term, such as "Latinx"? Several of our interviewees signaled that this kind of change is acceptable as long as it is noted in the text or a note. Speaking from the perspective of a data journalist, Lena Groeger suggested that this strategy is similar to paraphrasing an interview and that the creator should explain the decision and the process used to arrive at it. Others noted that such change would not be acceptable because survey participants responded to these terms, which may carry particular connotations. In such cases, the original survey language should be used in tables or charts, while an explanation of why this is not the preferred way of describing such populations can be woven into the text.

FIGURE 06

Legend where the labels use language that is not inclusive of different groups.



Source: Recreated from Tableau dashboard.

## 07 ORDERING DATA PURPOSEFULLY

Often, little thought is given to how estimates in tables or bars in graphs are ordered other than how they appear in the raw data. As graph producers, however, we should take a more active role in choosing how to order and present data values for different groups.

There is a historical legacy from which this sort of ordering and implied racial hierarchy derives. Many of the largest demographic surveys conducted in the United States order race starting with “white” and “Black” as the first two options. The images in figure 7 are the questions about race from the decennial census conducted by

the US Bureau of the Census, the Survey of Consumer Finances from the Federal Reserve Board, the Panel Study of Income Dynamics from the University of Michigan, and the Census Bureau’s Current Population Survey. Each lists “white” as the first option and codes that group with a “1” in the data. No wonder data analysts use it as the default option and thus the norm.

Which group we choose to show as the first row in a table or the first bar in a graph can affect how readers perceive the relationship or hierarchy between groups. Always starting with “white” or “men” can make these

FIGURE 07

Examples of survey questions about race in four major US federal government surveys illustrating how “white” is presented as the first response option.

**9. What is Person 1's race?**  
Mark ☒ one or more boxes **AND** print origins.

☐ White – Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.

☐ Black or African Am. – Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.

☐ American Indian or Alaska Native – Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.

<input type="checkbox"/> Chinese	<input type="checkbox"/> Vietnamese	<input type="checkbox"/> Native Hawaiian
<input type="checkbox"/> Filipino	<input type="checkbox"/> Korean	<input type="checkbox"/> Samoan
<input type="checkbox"/> Asian Indian	<input type="checkbox"/> Japanese	<input type="checkbox"/> Chamorro
<input type="checkbox"/> Other Asian – Print, for example, Pakistani, Cambodian, Hmong, etc. <input type="text"/>	<input type="checkbox"/> Other Pacific Islander – Print, for example, Tongan, Fijian, Marshallese, etc. <input type="text"/>	

☐ Some other race – Print race or origin.

**IN PERSON VERSION:**  
(SHOW CARD 20)  
Please look at this card. Which of these categories do you feel best describe you: (white, black or African-American, Hispanic or Latino, Asian, American Indian or Alaska Native, Hawaiian Native or other Pacific Islander, or another race?)

**TELEPHONE VERSION:**  
Which of these categories do you feel best describe you: white, black or African-American, Hispanic or Latino, Asian, American Indian or Alaska Native, Hawaiian Native or other Pacific Islander, or another race?

**CODE ALL THAT APPLY: CODE RESPONSES IN THE ORDER THEY ARE GIVEN**

1. \*WHITE (INCLUDE MIDDLE EASTERN/ARAB WITH WHITE); Caucasian
2. \*BLACK/AFRICAN-AMERICAN
3. \*HISPANIC/LATINO
4. \*ASIAN
5. \*AMERICAN INDIAN/ALASKA NATIVE
6. \*NATIVE HAWAIIAN/PACIFIC ISLANDER
- 7. \*OTHER
0. Inap. (/no further responses)

\*\*\*\*\*  
FOR THE PUBLIC DATA SET:  
ONLY X6809 AND X6810 ARE INCLUDED.  
FOR X6809, CODES 4, 5, AND 6 ARE COMBINED WITH CODE -7.  
IF AN ADDITIONAL RESPONSE WAS GIVEN IN X6810-X6814, X6810 IS SET TO 1; OTHERWISE X6810 IS SET TO 5.  
\*\*\*\*\*

**RACE**

I am going to read you a list of five race categories. You may choose one or more races. For this survey, Hispanic origin is not a race. (Are/Is) (NAME/you) White; Black or African American; American Indian or Alaska Native; Asian; OR Native Hawaiian or Other Pacific Islander?

\*Do not probe unless response is Hispanic or a Hispanic origin  
Enter all that apply, separate with commas

1	White
2	Black or African American
3	American Indian or Alaska Native
4	Asian
5	Native Hawaiian or Other Pacific Islander
6	Other - DO NOT READ

L7. What is your race? Are you white, black, American Indian, Alaska Native, Asian, Native Hawaiian or Other Pacific Islander?

- ENTER all that apply
- PROBE: Any others?
- If R says something not listed in response options, record under [7. Other] and list response verbatim at "specify screen"

1. White	2. Black, African-American, or Negro
3. American Indian or Alaska Native	4. Asian
5. Native Hawaiian or Pacific Islander	7. Source: other race-specify

→ L708. Specify (Using 500)

Sources: Decennial census from the US Census Bureau (top-left); Survey of Consumer Finances from the Board of Governors of the Federal Reserve System (top-right); Panel Study of Income Dynamics from the University of Michigan (bottom-right); and Current Population Survey from the US Census Bureau (bottom-left).

groups appear as the default against which other groups should be compared, suggesting they're the most important populations. How we choose to order may also reflect who we view as the intended audience for our visualizations. Again, starting with "white" or "men" can make it seem as though those are the most important groups we are trying to communicate with. A few issues are worth considering to help guide the decision of how to order different racial and ethnic groups:

- Does your study focus on a particular community? If it does, that group should be presented first.
- Is there a particular argument or story you are trying to tell? If so, the order or presentation of results should reflect that argument.
- Is there a quantitative relationship that can guide

how the groups are ordered? Can they be sorted alphabetically or by population size, sample size (weighted or unweighted), or magnitude or effect of the results?

Interactivity is another potential way to address the problematic aspects of choosing an order for your charts or tables. Enabling users to select which groups they want to see in a chart by, for example, toggling the lines that appear in a line chart or selecting which race appears as the baseline in a stacked bar chart, can empower users to focus on the groups they are most interested in or customize which comparisons they want to make. Instead of imposing a static view of a data visualization, allowing interactivity can help users find and explore the stories most relevant to them.

## 08 CONSIDERING THE MISSING GROUPS

It is also important to acknowledge who is and is not included in our data and charts. Many charts on race and ethnicity show white people, Black people, and Hispanic/Latino people but not other racial or ethnic groups. Many charts also present data at the level of these broader racial groupings rather than at a more disaggregated level. Often this is because of data limitations, usually small sample sizes or no data at all. But even in such cases, how can researchers help organizations conducting surveys be more inclusive? How can we communicate to those organizations to help them conduct better surveys? Just because it may be harder to obtain data about certain groups doesn't mean we shouldn't still try to better understand their lives.

Groups whose data are and are not collected or shown often reflects who society deems as most important or valuable. An example of this can be found in cartography, where maps of the United States often do not display US territories, erasing the populations that live there (Dougherty and Ilyankou 2021). Not only do the citizens of those areas not have any political representation in Congress (or the right to vote for president), these areas also have majority-nonwhite populations. Is it a coincidence that roughly 3.5 million people are not represented in these visualizations of the US? As mentioned in section 5, omitting certain populations

from the data can have important ramifications for what products or public policies are created and who benefits from those products and policies.

We can think about missing data even more broadly than data we have but do not include in our charts, graphs, and diagrams. Artist and researcher Mimi Qn̄q̄ha's *The Library of Missing Datasets* project identifies "missing datasets," or entire categories of data and information that do not exist but perhaps should. Qn̄q̄ha (n.d.) writes, "That which we ignore reveals more than what we give our attention do ... Spots that we've left blank reveal our hidden social biases and indifferences."

There are instances, however, in which collecting data on marginalized groups can be harmful, especially when done without their consent. Doing so can risk exposing people with adverse consequences for their livelihoods and safety (see our earlier note about data confidentiality and security). Some groups can also be overrepresented in datasets. In some instances, this is a purposeful (and useful) factor in conducting a survey that seeks to better represent small groups and the variation within those groups. In other cases, however, such overrepresentation can be harmful, such as when parents with low incomes were overrepresented in a dataset that was used as an input to predict the risk of child abuse in Allegheny County, Pennsylvania (Eubanks 2018).

In thinking about missing groups, we refer to four specific situations: lumping or splitting groups; using nonbinary gender categories; using catch-all groups often labeled as “other”; and choosing to not include all groups.

### **LUMPERS AND SPLITTERS**

For researchers and analysts, a big issue when deciding which groups to include in an analysis or show in a data visualization is the underlying sample size. When there are “too few” observations, populations may be lumped together to make analysis more convenient.

Doing so, however, can have harmful effects on the communities that are lumped together. The UCLA Center for Health Policy Research conducts outreach and works with Asian American, Native Hawaiian, and Pacific Islander (NHPI) communities. Ninez Ponce told us that “It’s a disservice to NHPIs to aggregate to an [Asian American and Pacific Islander] group because the generally better stats of Asians hide the vulnerabilities of NHPIs. That actually could be harmful for the Native Hawaiian Pacific Islander community.”

As a relatively simple example of how aggregating racial groups can mask important variation, we looked at the 2019 poverty rate across 139 detailed race categories recorded in the Census Bureau’s American Community Survey, an annual survey that helps determine how more than \$675 billion in government funds are distributed each year (US Bureau of the Census, n.d.).

The official poverty rate in the United States stood at 12.3 percent in 2019 (Benson 2020). Within that overall estimate, different racial and ethnic groups vary significantly. The primary question in the American Community Survey concerning race consists of 15 separate checkboxes; for some of these, interviewees are able to print additional names or races (a separate question refers to Hispanic or Latino origin). Digging slightly deeper into the data presents a more nuanced picture of poverty in the United States.

The dots in figure 8 show our estimated poverty rates for all 139 detailed racial groups from the American Community Survey, along with the overall poverty rate for major racial groupings frequently used in analyses (Hispanic/Latino is not shown in our chart because it is asked in a separate question about ethnicity). Within

the “American Indian or Alaska Native” category, for example, poverty rates vary from 5.9 percent for the Aleut (an indigenous community primarily living in Alaska) to 36.9 percent for those who identify as part of the Sioux Native American tribe (living primarily in Nebraska, North Dakota, and South Dakota). Within the “Asian or Pacific Islander” category, which is no longer a specific category in the American Community Survey but is often used in data analysis, we find a range of poverty rates, from 4.5 percent for those who identify as both Chinese and Japanese to 27.8 percent for those who identify as Mongolian. Even within the Pacific Islander category, we find a range of poverty rates, from 6.3 percent for Fijians to 24.6 percent for people who identify as more than one Micronesian race. We could tell similar stories about the more than 30 American Indian and Alaska Native tribes identified in the data, an issue that has raised calls for groups to take greater ownership of their own data (Rodriguez-Lonebear 2016). (It is also worth noting that interviewees from the NHPI Data Policy Lab stated that there is broad consensus among Asian American and NHPI organizations to retire the phrase “Asian and Pacific Islander” to help support more inclusive and accurate assessments of people in the United States; see Ishisaka 2020.)

A main reason more analyses don’t include more racial groups is likely sample size. In the data we used here, the number of people in poverty ranged from around 100 people for some groups to tens of thousands for others. Sample size limitations are a function of multiple factors, including the actual size of the group, staffing, and funding as well as the active choices made by the surveying organization, survey funders, and society. Heather Krause, founder of We All Count, noted in our interview that small sample sizes for groups that are already underrepresented is a “choice on the part of the privileged, not limitations that are inherent in small populations.” Thus, the burden to do a better job collecting data that reflects the lives of all people should not fall on already-marginalized communities but on the surveying and research institutions, which should strive for better representation of such groups in the full dataset.



### MALE, FEMALE, IN BETWEEN, AND BEYOND

Sometimes data are not collected at all about certain populations. Many major federal surveys, for example, do not offer “nonbinary” or “transgender” as response options when asking about gender. In such situations, data consumers should make it clear to their reader what information they have and do not have in their data. Lauren Klein said she includes a note in her work when certain groups are not included in the data, such as nonbinary people: “Just by calling out the absence and by doing it explicitly, it gives a prompt to the reader who may be the next survey designer to be like, ‘Maybe I should think about nonbinary gender.’” The entire ecosystem of data producers, consumers, and communicators should work in tandem to be more inclusive and remind each other of how these different aspects can be improved.

In January 2021, Billy Jacobson (also known as Anna Lytical, @theanalytical on Twitter) asked followers to submit “the worst gender selection forms.” A selection of those submissions are shown in figure 9 and demonstrates the challenges (and sometimes absurd results) websites and groups conducting surveys have in asking about missing groups. Although we are primarily

focusing on racial equity in this paper, we raise this issue around identifying gender to highlight both the challenges of missing data and the intersectionality between different characteristics that can raise even greater challenges to people collecting, analyzing, and communicating data.

This concept of missing data raises two further questions: First, do data collectors always need to ask for every piece of information? There seems to be a perspective that it is better to collect all possible pieces of information just in case additional research will be conducted later on. But asking for all this information may not be necessary. In a July 2020 blog post in *Health Affairs*, Rhea Boyd and coauthors recommend that health researchers “define race during the experimental design, and specify the reason for its use in the study” (Boyd et al. 2020). It may not be necessary to include race, gender, or other demographic variables in every data collection effort, project, or visualization.

As a personal example, the children of one of this guide’s authors (Schwabish) were participating in a virtual movie event with other kids during the COVID-19 pandemic.

FIGURE 08

Poverty rates across different racial groups when disaggregating data from the American Community Survey reveal variation that is missing when metrics are only presented for overall groups.

### Variation in Poverty Rates



Source: Schwabish and Feng (2021).



**Just by calling out the absence and by doing it explicitly, it gives a prompt to the reader who may be the next survey designer to be like, "Maybe I should think about nonbinary gender."**

– LAUREN KLEIN

Kids would register online and go to a curbside pickup location for bags of popcorn and hot chocolate. The registration form included basic information, such as the student's grade and parent's names, but it also provided a space for gender, which had no bearing on the event (figure 10). Website developers and data collectors should ask themselves whether collecting this information is necessary.

Second, products and research authored by more diverse teams can contribute to creating better and more inclusive products and projects because of the variety of perspectives and lived experiences team members bring with them. Such diversity can help survey and product designers identify gaps in the way they are collecting data. Perhaps the most famous example of how designing for a specific group can help everyone are the small ramps built into the curbs of sidewalks. The ramps

were initially an effort by disability advocates to assist people in wheelchairs, but they also serve parents with strollers, travelers with luggage, and workers pushing carts. This "curb-cut" effect shows how designing for one group can help many others (Blackwell 2017). Similarly, collecting inclusive data and building inclusive tools and visualizations can make the experience better for all users.

### OTHERING THE OTHER

A topic of particular interest for many people collecting and working with data is how to treat the "other" category. There are two primary issues here: First, what is the right approach to analyzing and communicating about the "other" category? Second, can we use more inclusive language to talk about this group?

On the first issue, we need to first understand what the "other" category is actually measuring. Depending on the data, the "other" category may include people with a wide array of races, ethnicities, religions, and other characteristics. (This is similar to the data lumping issue mentioned earlier, except here we do not have data that captures specifics about the individual and cannot disaggregate this category further.) Although "other" can be an important option for survey respondents who do not see themselves reflected in the usual racial

FIGURE 09

Selection of gender drop-down menus with questionable response options.

The figure displays three different gender selection interfaces. The top interface is a drop-down menu for 'Gender' with options: Female, Male, N/A, Unknown, and Tax Entity. The middle interface is a form with fields for GENDER, BIRTHDATE, and LANGUAGE, with GENDER set to Female, Male, Female, and Human. The bottom interface is a form for 'Spouse's gender' with radio buttons for Female, Male, and Indeterminate.

Source: Coding Drag Queen Anna Lytical (@theanalytical). 2021. "send me the worst gender selection forms you've seen, I'll start," Twitter, January 13, 2021, 11:25 a.m., <https://twitter.com/theanalytical/status/1349392166716657664>.

categories, we should question whether grouping people with such varied characteristics and life experiences together is actually meaningful. Although including all observations in the data sample is important for the math and statistical analysis, it may not be particularly useful in terms of policy solutions, as Calvin Chang from the UCLA team noted:

In terms of whether that would be useful at all, in terms of the organizations that you would expect to use a data report. It won't be useful for them to know what that coefficient is, or being able to dive any deeper into that "other" category, if their community is in the "other" category. So I can see it being part of the calculations, but in terms of the data being presented, it's just simply not going to be of much use. I could never go to a policymaker as a member of the NHPI community and say, "We need resources because the 'other' category doesn't have resources."

Regarding the issue of more inclusive language, it is important to recognize that the term "other" literally others individuals, emphasizing how they differ from the norm. The word "other" itself can be perceived to have negative

FIGURE 10

Example form that asks for gender even though it is not relevant.

Checkout

Grade \*  
Choose one

Gender \*  
Choose one

Parent/Guardian 1 First Name \*  
[Text Input]

Parent/Guardian 1 Last Name \*  
[Text Input]

Parent/Guardian 1 Cell Phone \*  
[Text Input]

Register

**Movie Night**

Order summary

Delivery 1 x eTicket \$0.00

**Total \$0.00**

Source: Original creation by authors in likeness of actual web form.

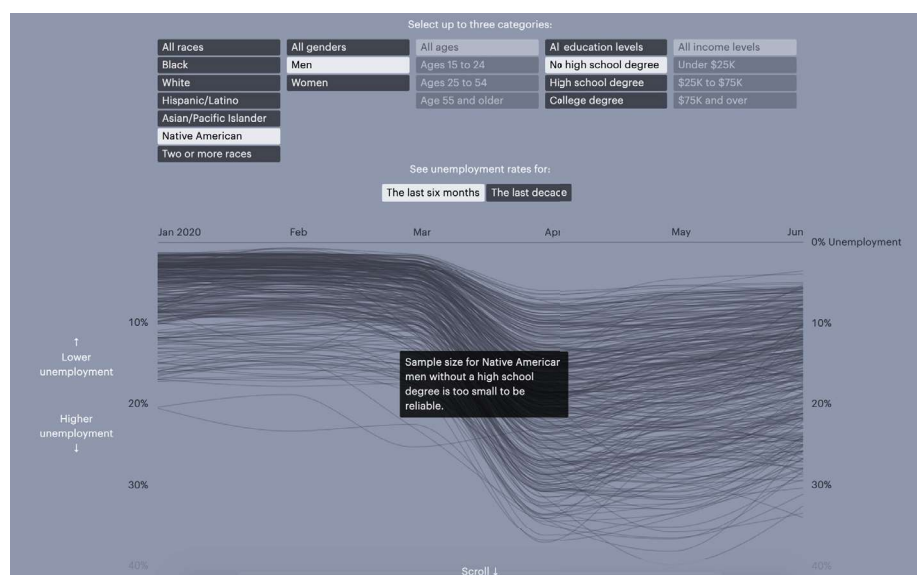
connotations. One meaning of "other" in the Merriam-Webster dictionary is "disturbingly or threateningly different," clearly evoking a negative or dangerous association.

Moving away from the word "other" is one step data communicators can take. Through our various discussions, we have identified six possible alternatives to "other":

1. Another race
2. Additional groups
3. All other self-descriptions

FIGURE 11

A message box explaining how sample size limitations did not allow for reliable estimates.



Source: "What Coronavirus Job Losses Reveal about Racism in America" (Groeger 2020).

4. People identifying as other or multiple races
5. Identity not listed
6. Identity not listed in the survey

Some of these terms are more verbose and may not fit as nicely in a table or under a bar in a bar chart, but they are more inclusive and avoid some of the issues noted.

As discussed in section 6, some may worry that these labels are not the exact words or phrases available to respondents in the survey. That is, the survey participant checked the box “other” because that was what was available. But we believe these alternatives are, in the general sense, still true and accurate; a separate note or footnote could be added to the graph, table, or text to make the change clear to readers.

### CHOOSING TO NOT INCLUDE ALL GROUPS

In cases where data were collected about a specific group but that group was not presented in the chart or included in the analysis, it may be worth listing in the notes section

all racial and ethnic groups included in the original dataset. This can both inform readers that data exist for these smaller groups (even though they may be limited) and make the chart-maker’s decision about groups they included in their visualization transparent.

ProPublica took another approach in its piece “What Coronavirus Job Losses Reveal about Racism in America” (Groeger 2020), a data tool that allows users to explore disparities in unemployment rates by race, gender, age, education, and income (figure 11). Some user selections would have produced sample sizes that are too small (such as Native American men without a high school degree). Rather than excluding those options from the tool and thus the population this combination of selections represents, ProPublica displayed a message over the chart explaining how the lack of sufficient data would result in estimates that were too imprecise to be reliable. This approach allowed ProPublica to be inclusive while not misinforming its readers.

## 09 USING COLORS WITH A RACIAL EQUITY AWARENESS

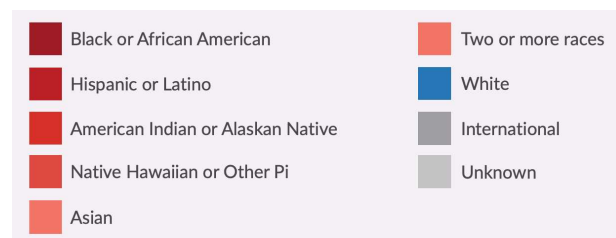
Good color palettes for data visualization should, at minimum, meet basic accessibility guidelines (WebAIM 2020) and offer sufficient contrast between colors for readers with vision difficulties. Going beyond accessibility, color choices should also avoid reinforcing gender or racial stereotypes, such as baby pink and baby blue to represent women and men or colors associated with skin tones or racial stereotypes (e.g., black to represent Black people, yellow to represent Asian people).

A data visualization that does not take a racially equitable approach to using colors is the Massachusetts Institute of Technology Office of the Provost’s “Diversity Dashboard” (MIT, n.d.), which enables users to explore the demographic characteristics of the school’s students, faculty, and staff. The June 2020 version of the dashboard used three distinct hues to represent nine racial and ethnic groups (figure 12). Six groups were represented with shades of red; the “international” and “unknown” groups were two shades of gray; and the “white” group was the only one in blue.

This design presents many problems. First, the shades of red that are used to represent the six groups of students of color create a visual divide that seems to pit students of color against white students. Second, a graduated color palette should not be used for categorical data. A graduated color palette shows greater or higher values in darker colors and smaller or lower values in lighter colors. Thus, the graduated palette here appears to suggest that “Black or African American” students are somehow

FIGURE 12

Legend showing a problematic color scheme applied to data on race and ethnicity.



Source: Recreated based on the June 2020 version of the *Diversity Dashboard* from the Massachusetts Institute of Technology, Office of the Provost.

“more” or “higher” than students who identify as “two or more races.” Instead, separate colors should be used for each of the six categories. Third, because gray colors tend to fade to the background (Schwabish 2021; Shirley 2020), this color choice diminishes international students and students whose race or ethnicity is unknown. All these design decisions create an effect where the “white” category moves to the foreground and is highlighted, as if that is the most important group and norm with which all other groups should be compared.

About a month after it was launched, the dashboard was updated but not dramatically improved (figure 13). Here, we see a blue color palette used for four student groups, a dark green representing Black or African American students, and a light purple for the “two or more races” category. White and international students now have similar orange/red colors, and the “unknown” category sits by itself to the side in grey.

Many of the same issues persist in this redesign. The use of graduated color palettes still suggests a hierarchy among races, and the dark green and light purple are close enough to the blue palette to still be perceived as being part of a singular color “ramp.” The similar orange/red colors for white and international students also seems to imply a relationship between these two groups.

It is not necessarily easy to pick out nine separate colors, but many free color tools are available for designers and developers. The palette in figure 14, for example, was chosen from the free ColorBrewer tool in less than a minute. Another possible approach would be to use a series of different graphs (the small multiples or panel chart approach) instead of putting all the data in a single view, which we discuss in more detail in section 14.

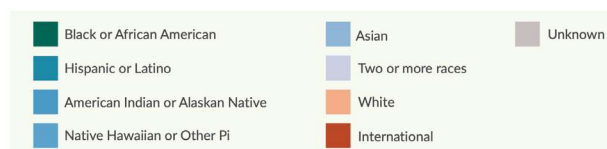
Another aspect of color to be aware of is the emotional connotations associated with certain hues. In Western

cultures, colors such as red can be perceived as threatening or aggressive and therefore can paint the population we’re visualizing in a negative light. In a map of participation in employment services for those receiving welfare benefits in Minnesota, participants were represented as small red dots. When the participants who were being visualized in the map were shown that map, they felt that the design presented them as a “threat” and made them look like an “infestation” (Evergreen, n.d.). The implication of this visual approach is that the problem lies with the people rather than the legacy of racist institutions and policies. Again, it is important to ask, “If I were one of the data points on this visualization, would I feel offended?”

In sum, as data visualization producers, we need to be aware of how our use of colors, words, and categorizations can perpetuate or exacerbate inequities and stereotypes.

FIGURE 13

Legend with the color scheme modified (though still problematic).



Source: Recreated based on the July 2020 version of the *Diversity Dashboard* from the Massachusetts Institute of Technology, Office of the Provost.

FIGURE 14

Example color palette that avoids gradients and hierarchies.



Source: Created with the ColorBrewer online tool.

## 10 USING ICONS AND SHAPES WITH A RACIAL EQUITY AWARENESS

We want to be careful and thoughtful when using icons in any data visualization. When showing groups of people, we should consider a mix of genders, races, ethnicities, ages, and other characteristics. When using icons, we should consider to whom we are presenting

our results, anticipate how our icons might be perceived, and ensure they do not perpetuate harmful and offensive stereotypes. We need to be conscious of how certain icons may not correspond to the content, such as an icon of a baby in a chart about child mortality.

Mis- or underrepresentation of certain groups in imagery and iconography can fail to take a racial or gender equity awareness perspective toward how we visualize our data. A 2018 study by the Pew Research Center found that “men are overrepresented in online image search results across a majority of jobs examined” and that “women appear lower than men in such search results for many jobs.” Similarly, scholar Safiya Noble (2018), in her book *Algorithms of Oppression*, records a racist and sexist culture in our online searches and other actions on the internet. These disparities continue to this day; the top image in figure 15 shows Google search results for the phrase “nurse icon,” and the image on the bottom shows results for the phrase “boss icon.” Notice how nearly all the images on the top we might imagine as feminine and the images on the bottom we might traditionally consider masculine.

When choosing icons or imagery to represent racial or ethnic groups, avoid images that depict or reinforce stereotypes associated with those groups, such as African Americans in poverty or American Indians in traditional headdresses on horseback (Kim 2019). Pictures also should not show people in situations where they appear as helpless victims (so-called “poverty porn”; Dortonne 2016) or reinforce power hierarchies (e.g., a white male supervisor with a person of color as a subordinate). Instead, images should represent people as active and empowered and reinforce their dignity, agency, and humanity. (A separate but related issue is how to use emojis with different skin tones, which has spurred all sorts of discussions and debates about whether the original, default yellow skin tone represents a white default; see McGill 2016.) And when possible, try to solicit feedback about the visuals used to make sure they are not offensive and to identify problematic cultural associations that were missed.

Another factor to consider when using icons or shapes that resemble people and communities is when they might imply something untrue about the data. The 2019 ProPublica story “What Happened to All the Jobs Trump Promised?” (Porat, Groeger, and Arnsdorf 2019) used

icons of people to represent the number of jobs President Trump claimed to have created (figure 16). In designing this story, Lena Groeger and her colleagues debated whether it was appropriate to include an icon with a person in a wheelchair. Including such an icon could be seen as representing people with disabilities and making them more visible. But readers might also interpret the icon to mean that the people in those jobs were disabled when that might not be true. In the end, Groeger and her colleagues decided not to include the icon of the person in a wheelchair. Ultimately, one should exercise caution when using highly specific icons.

FIGURE 15

Image search results showing problematic stereotypes and biases.

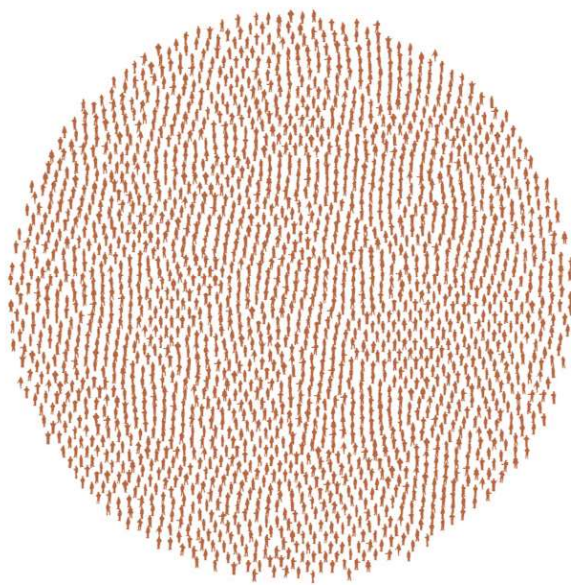


Source: 2021 Google image search of “nurse icon” and “boss icon.”



FIGURE 16

A chart that uses icons representing people.



(Each  represents about 4,000 jobs)

Since the election, Trump has made 35 claims that companies would create 8.9 million jobs in the U.S. thanks to his policies and actions. Some of the new jobs he touted failed to materialize. Many of the new jobs he took credit for were, in fact, planned before he took office.

Source: "What Happened to All the Jobs Trump Promised?" (Porat, Groeger, and Arnsdorf 2019).

## 11 EMBRACING CONTEXT AND COMPLEXITY

### If I only saw this chart on Twitter, would I draw the correct conclusion?

– KIM BUI

As mentioned in section 6, it is important that data visualizations acknowledge the context and history surrounding the topic being shown. Structural racism, historic discrimination, other barriers and inequities, and the mechanisms by which they may be operating should be woven into the framing of the research and the chart (Schwabish and Kijakazi 2021). Visualizations should not let the data “speak for themselves”: data are not neutral or objective, so data visualizations are not neutral either (as we saw with the example of the HOLC redlining map in the introduction). When data visualizations are not critically examined and thoughtfully designed within a DEI framework, they tend to express the viewpoint of the dominant group in society while, as Catherine

D'Ignazio and Lauren Klein (2020, 76) write in their book *Data Feminism*, “mask[ing] the people, the methods, the questions, and the messiness that lies behind clean lines and geometric shapes.” When weaving in context, data communicators should have diverse references and citations, particularly from scholars of color or scholars who are members of the community of focus, and they should lift up the voices of the communities they worked with, highlighting lived experience where possible.

Data communicators and visualizers also should not shy away from embracing complexity and nuance in their visuals. Even though simple charts and the idea of having a high “data-ink ratio” (Tufte 1985) are often lauded as goals in data visualization, too much simplicity can strip away necessary context and the inherent complexity of many social issues that cannot be boiled down to a simple bar chart. As Kim Bui put it, “If I only saw this chart on Twitter, would I draw the correct conclusion?”



Data visualizations should use more complex designs if that would more accurately reflect and promote a better understanding of the topic being shown.

Complex designs can also foster better audience connection and engagement with the data. In her articulation of “Data Humanism,” information designer Georgia Lupi argues that dense, custom-designed data visualizations presenting multiple layers of information can encourage more careful reading, better personal connection, and a deeper understanding as readers are offered various paths to explore the data. Using a more personal approach to visualizing data, one that incorporates qualitative information, expresses subjectivity, and acknowledges uncertainty can “unlock [data’s] profound nature and shed light on its real meaning for ourselves and others” (Lupi 2017). Data visualizations should be clear and informative while not oversimplifying the complex world we live in.

It is worth noting that not all data needs to be visualized; not every graph needs to be created. Data communicators should consider whether their instinct to create a graph, chart, or diagram is warranted. Sometimes a simple number or set of numbers can be more effective than a graph. And sometimes—perhaps many times—data visualizations are unable to convey the complexity of the

underlying data or the bias inherent in the data or the visuals. As Sarah Williams (2020, 184) writes in her book, *Data Action*, “Sharing data through visual interpretation might make underlying policies more transparent, but all data visualizations hold the bias of the creator.” It is important to remember that data are a reflection of the lives of real people, not just a sterile abstraction.

Further, charts simply may not be as effective for forming an empathetic connection to the subject as other visual or storytelling methods (Correll 2019). In a story on families with children who have a fatal disease, Sarah Slobin and her colleagues at the *Wall Street Journal* at first tried to visualize data on an index that tracked the progression of the disease (Marcus 2013; Slobin 2014). After struggling to find a compelling way to visualize these data and the children affected by it, Slobin and her team realized that charts were not necessary after all and that simply showing photos of children diagnosed with these fatal diseases was sufficient, perhaps even more compelling: “When you looked at the kids—really looked at them, even though it was hard to look at them—you understood what the data was capturing, how the disease progressed, what it wrought. And you saw these beautiful little people and you understood what the Severity Index was all about without a chart or a visualization or an explainer” (Slobin 2014).

## 12 BUILDING DIVERSE AND INCLUSIVE DATA COMMUNICATION TEAMS AND ORGANIZATIONS

Across most of our interviews, we heard that diverse teams, in terms of demographics, life experiences, skills, and subject matter expertise, are critical to creating communication products that apply a racial equity lens and fight other forms of oppression. Diverse teams can help identify biases and make connections between different fields of study whose relevance may not be evident at first glance. They can also better reflect the demographics of the populations they wish to study. However, having diverse teams is not in itself sufficient. The internal culture of the organization should also value, promote, and embody DEI so that the communication products that are created uphold such values. The concept of building teams rather than leaving the

responsibility in the hands of individuals is a primary theme that emerged from our conversations.

We can distill our conversations around building diverse and inclusive organizations into four main areas:

### 1. A ground-up approach and buy-in from leadership go hand in hand.

In many organizations, change first occurs because an individual or small group of people see an opportunity to improve or evolve. In his work at the *Washington Post*, Aaron Williams saw this firsthand: “I think that there were reporters there first who were like, ‘I’m going to start doing this in my work,’ and then editors who got behind their reporters and then it started bubbling up.” Kim Leary echoed this point, saying

that individuals can ask questions such as “Why not?” and “How come?” challenging existing institutional norms and bringing creativity and innovation needed for fresh thinking. Though efforts may start at the individual level, buy-in and support are eventually needed from leaders and management to ensure that change is implemented throughout the institution.

## **2. Processes for incorporating equity and inclusiveness need to be codified for change to happen.**

Working toward a goal of equity and inclusiveness is fine at the conceptual level, but concrete, actionable guidelines need to come out of such discussions for organizations to successfully implement changes to their culture and the way they work. Guidelines must be codified (through documentation, checklists, guidelines, requirements, and testing) and made available to all members of the organization so everyone is following a standardized process.

## **3. Guidelines should be actionable and instill accountability.**

Ayesha McAdams-Mahmoud from Salesforce told us about her experiences building toolkits and resources for product developers to implement in order to build accessible and equitable tools. Those guidelines needed to include concrete steps that could be taken and “baked into the product development process” for them to be useful to developers. Those toolkits included measures of accountability, which are not necessarily punitive actions but rather guidelines and metrics that imbued a sense of responsibility. It is important at this step for organizations to also examine the official and unofficial practices and norms that incentivize or reward people for doing (or not doing) certain things.

## **4. Updating and reevaluating norms should have a regular cadence.**

Ongoing checks, team conversations, and retrospective analyses can all be ways to ensure products and projects are taking a racial equity lens and instituting organizational guidelines and norms. Guidelines and norms themselves should be updated and reevaluated periodically as well (Chartoff 2018).

Cultural shifts of this magnitude need to happen at both the organizational level and the individual and interpersonal levels. McAdams-Mahmoud and others talked about the importance of working with members

of the community when it comes to external data communication, but that theme also emerged when we discussed creating cultural change within an organization. From McAdams-Mahmoud’s perspective, involving individuals at all levels helps drive home a sense of ownership among the entire organization: “[We want to] make sure we’re bringing people along before we institute these processes. Our hope is that it’s not a surprise to anybody once these checklists come down the barrel because we’ve already said, ‘Hey this is important. We’ve had conversations about it.’”

The question of making cultural change then becomes, “How do we persuade people to do something, especially something that could be viewed as burdensome or extra work?” How to effect cultural change, change habits, and help businesses evolve is beyond the scope of this work, but three examples of major public campaigns from the past 50 years show how it can be done:

- Antismoking campaigns
- “Don’t Mess with Texas,” a campaign to reduce litter in the state
- “Click It or Ticket” and other efforts to encourage Americans to wear seatbelts

Ongoing antismoking campaigns (Durkin, Biener, and Wakefield 2009), along with changes in smoke-free laws and changing prices for cigarettes (American Heart Association 2018), have contributed to significant declines in the percentage of people who smoke. The original “Don’t Mess with Texas” slogan was created by the Texas Department of Transportation to reduce litter and used TV commercials featuring famous Texans as well as bumper stickers and other merchandise to spread the message. And it worked: between 1986 and 1990, road litter declined 72 percent (Nodjimbadem 2017). Finally, the “Click It or Ticket” campaign, which started in 1993 in North Carolina and has since expanded nationwide, is credited with increasing seat-belt usage and saving thousands of lives (Solomon, Compto, and Preusser 2004; Tison et al. 2010). When it comes to persuading individuals and organizations to change, varied approaches are needed: both “carrots” that help individuals see the benefits of such change and “sticks” that increase the cost of not adopting these changes.

## 13 THE FEEDBACK LOOP

In section 4, we noted the importance of working closely with and empowering the people and communities our research focuses on. This includes using a variety of research methods—such as interviews, focus groups, and long-form surveys—for a given body of work. Another approach is to solicit feedback from communities and audiences about our work. This can allow researchers and communicators to engage with audiences, understand how their work is being used and what they may have missed, and continue to develop and build relationships with communities. Across our discussions, we saw various approaches that different organizations took to soliciting feedback.

**ProPublica.** The ProPublica newsroom has an engagement team that works with reporters to communicate with readers, commentors, survey respondents, and more. These team members will sometimes be responsible for fielding calls and comments and will work directly with reporters on specific stories. ProPublica also often includes a form at the end of certain stories that provides readers with an opportunity to share their experiences or provide more information on how an event or policy affected them, as well as whether they are involved in the issue in some way.

**UCLA Center for Health Policy Research.** The UCLA team has several internal and external ways to receive feedback on data projects. Internally, they have a significant review process: team members who identify as different races and ethnicities review each product they produce, including infographics and other standalone visual content. Externally, the UCLA team has invested time in building relationships with the people and communities that are in their content focus areas. For example, when developing a survey for the Native Hawaiian and Pacific Islander community, the team created posters and flyers to generate awareness of the survey and asked members of the community to review the products. They also highlighted that the diversity of their own team helps them build connections and develop relationships with community groups who can then see themselves reflected in the research team. In our interview, they noted they have found it essential

to implement community-based participatory research principles (discussed also in section 9), which emphasizes incorporating both input and feedback loops from community organizations and leaders from the initial planning and design phase through implementation and dissemination.

**Evan Peck.** Evan Peck and his team at Bucknell University applied the concept of “reciprocal research” to how they conduct their projects. Reciprocal research (Hanrahan et al 2020; also discussed in section 9) is the idea of trying to create immediate, concrete benefits project participants can receive while they are participating in the study rather than waiting until the project is over before seeing how such work helps them. As Peck states, “You aren’t just getting data for your sake and something that might benefit the community back in 15 years, but you start thinking a little bit more short-term or how can you provide benefit back to the community as you’re collecting this data ... it’s how to be more deliberate in making sure we’re not unintentionally exploiting these people’s experiences in ways that don’t benefit them.” In applying such an approach, researchers may need to be flexible and responsive with their plans, pivoting to a new approach when they realize their original ideas do not meet the needs of the study’s participants and community.

---

**You start thinking a little bit more short-term or how can you provide benefit back to the community as you're collecting this data ... it's how to be more deliberate in making sure we're not unintentionally exploiting these people's experiences in ways that don't benefit them.**

– EVAN PECK

---

**Institute of Governmental Studies at the University of California, Berkeley.** G. Cristina Mora described what we perceive to be a more common approach in the academic fields of having a faculty advisory board who is responsible for reviewing certain projects for equity

and inclusion. As Mora described, the Berkeley advisory board consists of about 20 faculty across campus and is diverse in race, ethnicity, and gender. Mora noted that participating in these kinds of advisory positions can be an extra burden on faculty members' time, so compensation and consideration of that time is central: "You have to compensate them. Faculty of color are so taxed already ... We get most of the graduate students of color. We sit on almost every diversity committee that

exists on campus, and we still have to do our own work ... so you [have to] have creative ways of making this not be an additional tax on them and their time."

There are many ways to solicit and receive feedback, but the overarching themes in these and other conversations is to (1) build relationships with people and communities over time and (2) value people's time and compensate them accordingly.

## 14 THE GATEKEEPERS

Most data communicators, especially researchers and analysts, rely on funders for financial support to perform their work and on peer-reviewed publications to communicate the findings of their research. Thus, gatekeepers—funders, project officers, editors, peer reviewers, federal agencies, and others—can play an important role in enforcing a racial equity lens in the work they fund and publish because of the amount of influence they have on researchers' careers. Here, we discuss how three of these gatekeepers can help promote racial equity across the data ecosystem.

### EDITORS AND PEER REVIEWERS

Attributing findings to structural racism is not yet common in scholarly writing for many fields. Regarding economics, Howard University professor William Spriggs wrote, "The fact that far too many economists blindly agree that negative attributes correlate to being African American and cannot see that relationship to police officers assuming all Black men are criminals is stupefying. The fact that a discipline that prides itself on being objective and looking for data to test hypotheses fails to see how negative attributes do not correlate with being African American is a constant irritant for Black economists" (Spriggs 2020). In a July 2020 blog post in *Health Affairs*, Rhea Boyd and coauthors wrote that "despite racism's alarming impact on health and the wealth of scholarship that outlines its ill effects, preeminent scholars and the journals that publish them, including *Health Affairs*, routinely fail to interrogate racism as a critical driver of racial health inequities" (Boyd et al. 2020). Researchers and scholars need to do a better job of examining how racism impacts their findings, and

journal editors and reviewers can play an important role in demanding this happens (Schwabish and Kijakazi 2021).

Between February and May 2020, over 807,000 manuscripts were submitted to more than 2,300 journals on the Elsevier publishing platform (Squazzoni et al. 2020). Each of these manuscripts was presumably reviewed by an editor and sent to experts in the field to review and comment. If editors and reviewers asked authors specific questions about racial equity for each manuscript they review, we would likely see a shift in how race is analyzed, visualized, and written about in academic publishing.

Journals could also start to require "positionality statements" as part of any manuscript submission or publication. A positionality statement describes an individual's social and political context that helps shape their identity in terms of race, class, gender, sexuality, and ability status. Such positionality statements—in short, a summary of the author's lived experiences—may not only highlight existing gaps in the composition of research teams (e.g., an all-white, all-male research team writing about Black women's fertility), but it might also help the research team become more attuned to its own possible biases and perspectives as it relates to their work. (Obviously, journals will need to carefully consider how such statements would work with blind reviews and author privacy.) As Heather Krause noted, "You're not going to get actual empathy if there isn't first some examination of privilege." Positionality statements should examine not only individual team members' identities but also the identities of their team as a whole and that

of the institution they work in, as well as the implications of those identities. Such identities necessarily affect the biases we hold, which in turn influence the decisions made during a research project, such as the questions that are asked, how data are gathered, how findings are interpreted, and who the main audience is for the work. Journals are increasingly requiring financial and data disclosure documents as part of the publication process; positionality statements may be an appropriate next step.

### **GOVERNMENT AGENCIES, REGULATIONS, AND DIRECTIVES**

In addition to being funders of research, federal governments have significant power and influence on how surveys are designed and thus the data that are collected, which in turn affects how research is conducted. In 1977, the Office of Management and Budget adopted Directive Number 15 (OMB 1977), which directed federal statistical agencies to collect and report data on race and ethnicity across four different race categories (American Indian or Alaska Native, Asian or Pacific Islander, Black, and white) and one ethnic category (Hispanic origin). In 1997, OMB revised that directive to define five minimum categories for race (American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Black, and white) and changed the ethnic label designation from “Hispanic” to “Hispanic or Latino” (OMB 1997). Federal agencies follow these directives, and this directly influences how other firms and companies collect their own data. If the OMB were to revisit these categories that were created more than 20 years ago, other agencies, firms, and organizations would follow.

Another agency with significant influence in how racial and ethnic data are reported in the United States is the Bureau of the Census. Charged with collecting and publishing some of the most widely used demographic and economic data in the country (including the constitutionally mandated decennial census), the Bureau has what G. Cristina Mora described as “symbolic power” around defining race and ethnicity:

It's what we call symbolic power. It's like the cultural, symbolic power of the state. And to define who we are lies often really largely in the Census Bureau

in particular....The whole idea that even South Asians right now would see themselves, however tenuously, as part of the Asian-American category is through the power of the Census Bureau. Nobody else ever called them part of having anything to do with East Asians. And now we see it 30 years later, it's pretty strong and a tangible identity that's connected to it. It's not perfect, but that's the power of the Census.

Were the Census Bureau to, for example, revise how it asks about race and ethnicity—that is, how questions about race (e.g., white or Black) are separate from questions about ethnicity (e.g., Hispanic or Latino)—and how it publishes and codes those responses, researchers and analysts would change their approach to discussing these groups as well.

### **FUNDERS**

Many organizations and individual researchers are responsible for raising funds to support their work. Just as researchers should be held responsible for taking a racial equity lens to their work, funders should be held to similar account in the funds they disburse and their requests for analysis. In some cases, funders may request data in formats and categories that are short sighted or that do not align with best practices. As an example, in summer 2020, we (the authors) were asked to create a series of charts for a membership organization to highlight the challenges their readers were facing during the COVID-19 pandemic. We insisted that data around race and ethnicity play a central part in the data-driven project even though it was not explicitly asked for. Without that focus, we argued, the analysis would fail to acknowledge the adverse impacts people and communities of color were facing during the pandemic.

In other cases, funders may not recognize the need for overarching changes to their funding models or project processes. In the early part of summer 2020, the Rockefeller, W.K. Kellogg, and Walmart Foundations issued a wide-ranging request for proposals (RFP) to assess the impacts of COVID-19 pandemic responses across the food system. In response to the quick-



turnaround RFP, 17 Black, Indigenous, and other people of color (BIPOC)–led groups working in the food and agriculture sectors called for the RFP to be canceled and to “move in concert with BIPOC-led orgs doing work in BIPOC communities” (HEAL Food Alliance 2020). They argued the call reinforced existing inequalities in the food and agricultural justice sector by offering funding to “white-led organizations to do service work in BIPOC communities, or to fund a white-led organization with an established funder relationship to subgrant to an under-resourced BIPOC-led organization.” This, combined with the short time to submit project proposals (the RFP was issued on June 8, 2020, and proposals were due on July

1, 2020) that would require multiple teams and hundreds of thousands of dollars (if not more than a million) to put together, effectively cut smaller, BIPOC-led organizations from participation. Other organizations elevated these critiques, and the funders ultimately pulled the RFP. The organizations suggested several steps the funders could take in the future, including cultivating real partnerships with BIPOC communities and BIPOC-led organizations; targeting funding to smaller, BIPOC-led organizations; and investing in multiyear grants using a participatory grantmaking model.

## 15 CASE STUDY FROM THE URBAN INSTITUTE

We now walk through an example when the Urban Institute data visualization team sought to more consciously and critically apply a racial equity lens to a data visualization project. “Tracking COVID-19’s Effects by Race and Ethnicity” (Urban Institute, n.d.), which was published in summer 2020, is a data tool that enables users to visualize the disparate impact of the COVID-19 pandemic by race and ethnicity for various metrics. Initially, the central graphic was a series of line charts that showed how specific metrics had changed over time for each of five racial or ethnic groups. These charts displayed all the races together, with each racial group represented by its own line, a common approach for visualizing this type of data. The image in figure 17 shows, by racial and ethnic groups, the national share of adults in households with children enrolled in public or private school where at least one child had class canceled because of the pandemic between April and May 2020.

Given this project’s focus on racial and ethnic disparities, however, we questioned whether our initial design for visualizing this data truly reflected a racial-equity-aware approach. Though it is common to present data disaggregated by race or ethnicity with all groups on the same chart, this can have negative effects. As consultant Pieta Blakely writes, showing all groups together “encourages [the reader] to compare each of the groups to the highest performing groups,” which can lead to a “deficit-based perspective” that focuses attention

on what low-performing groups are lacking compared with the high performers (Blakely 2019). In addition to portraying members of groups that have been harder hit by the pandemic in a negative light, readers may neglect the needs of individuals within groups that appear to be, on average, doing “well,” even though members from all races and ethnicities have been negatively affected by the pandemic.

Some counter that this deficit framing perspective is not always applicable depending on how much context regarding the issue being visualized is offered. The UCLA team raised an important point that showing all racial and ethnic groups on the same chart might not be interpreted with a deficit-based lens if sufficient context on the historical inequities and discrimination that have led to these differences between groups is presented. Members of those groups may not necessarily view the chart as deficit based because they understand the gaps are the result of historical discrimination. In the case of our project, however, we felt we did not go into such historical factors in sufficient detail to mitigate the possibility of a deficit-based interpretation.

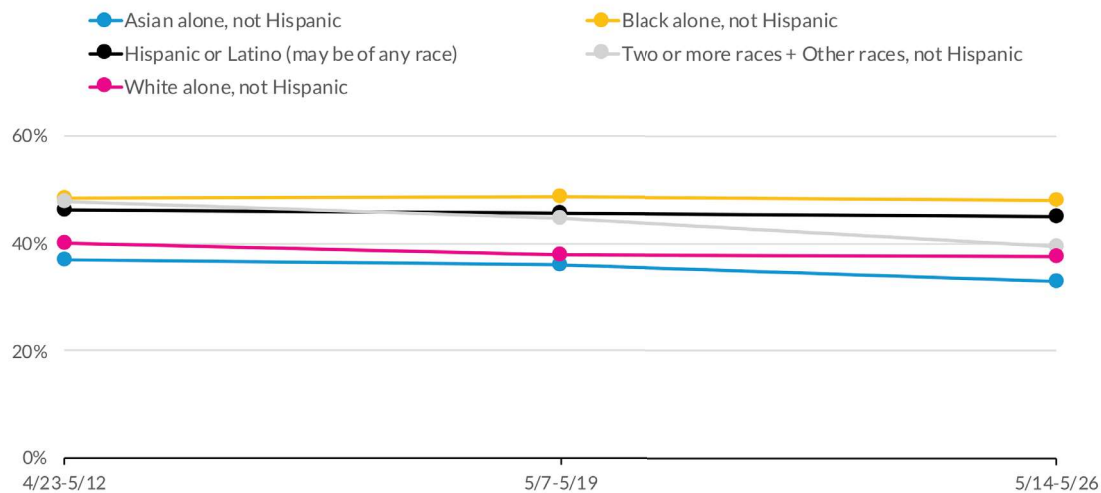
Another drawback of visualizing data in this manner is that such charts can also make it seem that the racial group least impacted by the pandemic is what the other racial groups should seek to match. But we should ask ourselves why, for example, is 12 percent of households experiencing food insufficiency acceptable just because that’s the share of white households in that situation at a particular point



FIGURE 17

Image from an early draft of the Urban Institute's "Tracking COVID-19's Effects by Race and Ethnicity" project.

### Share of adults in households with children enrolled in public or private school where at least one child had class cancelled because of COVID-19



Source: Urban Institute (n.d.).

during the pandemic? The real goal to strive toward is eliminating food insufficiency altogether, not just bringing all racial groups to the same level as the population that has been least negatively impacted. As such, it is worth thinking about how the way we visualize data might reinforce or challenge the way we frame these types of social challenges.

In the end, we shifted our approach to use multiple smaller charts (known as panel charts, trellis charts, or small multiples) rather than plotting all groups on the same visual (figure 18). We felt that by showing each race and ethnicity individually, this design might better encourage readers to think about the specific needs and challenges facing each group. Having six charts to work with (one for each demographic group and one showing all groups) also allowed us to add the relevant location-specific average as a consistent benchmark.

In ordering the six charts, we consciously sorted the groups alphabetically rather than by data value. Because this is an interactive data tool where users can select different metrics, sorting the charts alphabetically

maintains the same order across all views rather than having the graphs shift around, which can be disruptive to the user experience.

Finally, rather than showing estimates as lines, we displayed each data point as a dot surrounded by confidence interval blocks to reflect the uncertainty around our estimates caused by statistical and data collection issues. We used opacity to signal when differences were or were not statistically significant (i.e., more transparent blocks and dots indicate that the difference between the group-specific and the statewide estimates was not statistically significant; more opaque blocks and dots indicate the difference was statistically significant). This was important to show because often the data collected did not allow us to make precise estimates, and we wanted to be transparent about these limitations so readers would not incorrectly interpret what these charts are showing or falsely conclude there was or was not a racial or ethnic disparity when that may not have been the case.

FIGURE 18

Image of the final design of the Urban Institute's "Tracking COVID-19's Effects by Race and Ethnicity" project.

Share of adults in households where at least one person has lost employment income since March 13

● National average ● State or metro area average ● Race/ethnicity average  
 ■ Statistically significant difference ■ No statistically significant difference



Source: Urban Institute (n.d.).

## 16 LOOKING FORWARD

---

We have laid out many recommendations and guidelines that people and organizations should consider when taking a racially equitable view of their data and data visualizations. The issues we have highlighted here are by no means exhaustive—we have not explored issues relating to, for example, data visualization interactivity, accessibility, database structure and management, cultural differences, data privacy and security, and myriad other issues people working with data should consider daily. Nor are our recommendations ironclad rules. We view the principles outlined here as a starting point for people to think more carefully and critically about how to embrace equity and inclusion throughout the data pipeline of creation, analysis, and communication. Rather than dictating the right or wrong approach, we urge analysts and data communicators to be aware of the decisions they have made and know why they made those decisions.

Many issues we have highlighted here and that our interviewees brought to our attention will change and evolve with society and technology. And experts in data science, data communication, and similar fields are in positions to improve how people view and understand race, ethnicity, and other aspects of identity. Moreover, those views will in turn affect how different policies are created and implemented or products designed and built. Doing so through an equitable and inclusive lens will ultimately create a better society for everyone.

## 17 ACKNOWLEDGMENTS

The authors thank Greg Acs, Shena Ashley, Kreg Steven Brown, Michael Correll, Heather Hahn, Allen R. Hillery, Billy Jacobson (a.k.a. Anna Lytical), Kilolo Kijakazi, Marla McDaniel, Alexandra Tamaro, and Aaron Williams for their thoughtful comments and feedback. We also thank the people we interviewed during this project for their invaluable insights and experiences: Kim Bui, Kim Campbell, Richard Calvin Chang, Catherine D'Ignazio, Arthur Gales, Samir Gambhir, Lena Groeger, Lauren Klein, Heather Krause, Kimberlyn Leary, Tiffany Lopes, Ayesha McAdams-Mahmoud, Tim Meko, G. Cristina Mora, Evan Peck, Ninez Ponce, Peter Mattingly, Jacob Rosalez, and Aaron Williams.

### THE PEOPLE WE INTERVIEWED

We are especially thankful to the people we interviewed during this project for their invaluable insights and experiences.



KIM CAMPBELL



RICHARD  
CALVIN CHANG



CATHERINE  
D'IGNAZIO



SAMIR GAMBHIR



LENA GROEGER



LAUREN KLEIN



HEATHER KRAUSE



KIMBERLYN LEARY



TIM MEKO



G. CRISTINA MORA



EVAN PECK



NINEZ PONCE



AARON WILLIAMS

### NOT PICTURED ABOVE

Arthur Gales

Ayesha McAdams-Mahmoud

Jacob Rosalez

Kim Bui

Peter Mattingly

Tiffany Lopes

## 18 REFERENCES

- American Heart Association. 2018. "Smoking in America: Why More Americans Are Kicking the Habit." American Heart Association, August 30, 2018. <https://www.heart.org/en/news/2018/08/29/smoking-in-america-why-more-americans-are-kicking-the-habit>.
- Bartley, Lawrence, Lisette Bamenga, Adria Watson, Rahsaan Thomas, and Wilbert L. Cooper. 2021. "The Language Project." New York: The Marshall Project. <https://www.themarshallproject.org/2021/04/12/the-language-project>.
- Benson, Craig. "Poverty: 2018 and 2019, American Community Survey Briefs." Suitland–Silver Hill, MD: US Bureau of the Census. <https://www.census.gov/content/dam/Census/library/publications/2020/acs/acsbr20-04.pdf>.
- Betzendahl, Lindsay. 2020. "Racial Inequality in the United States." Tableau public data visualization, June 24. <https://public.tableau.com/en-us/gallery/racial-inequality-united-states?tab=viz-of-the-day>.
- Blackwell, Angela Glover. 2017. "The Curb-Cut Effect." *Stanford Social Innovation Review*, Winter. [https://ssir.org/articles/entry/the\\_curb\\_cut\\_effect](https://ssir.org/articles/entry/the_curb_cut_effect).
- Blakely, Pieta. 2019. "Presenting Data for a Targeted Universalist Approach." *PietaBlakely.com* (blog), March 6, 2019. <http://blog.pietablakely.com/presenting-data-for-a-targeted-universalist-approach>.
- Blas, Terry. 2019. "'Latinx' Is Growing in Popularity. I Made a Comic to Help You Understand Why." *Vox*, October 23, 2019. <https://www.vox.com/the-highlight/2019/10/15/20914347/latin-latina-latino-latinx-means>.
- Borkin, Michelle A., Zoya Bylinskii, Nam Wook Kim, Constance May Bainbridge, Chelsea S. Yeh, Daniel Borkin, Hanspeter Pfister, and Aude Oliva. 2015. "Beyond Memorability: Visualization Recognition and Recall." *IEEE Transactions on Visualization and Computer Graphics* 22 (1): 519–28.
- Bowen, Claire, Aaron R. Williams, and Ajjit Narayanan. 2021. "To Advance Racial Equity, Releasing Disaggregated Data while Protecting Privacy Will Be Key." *Urban Wire*, March 2, 2021. <https://www.urban.org/urban-wire/advance-racial-equity-releasing-disaggregated-data-while-protecting-privacy-will-be-key>.
- Bowen, Claire. 2020. "Will the Census's Data Privacy Efforts Erase Rural America?" *Urban Wire*, March 4, 2020. <https://www.urban.org/urban-wire/will-census-data-privacy-efforts-erase-rural-america>.
- Boy, Jeremy, Anshul Vikram Pandey, John Emerson, Margaret Satterthwaite, Oded Nov, and Enrico Bertini. 2017. "Showing People behind Data: Does Anthropomorphizing Visualizations Elicit More Empathy for Human Rights Data?" In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, by Association for Computing Machinery, 5462–74. New York: Association for Computing Machinery.
- Boyd, Rhea W., Edwin G. Lindo, Lachelle D. Weeks, and Monica R. McLemore. 2020. "On Racism: A New Standard for Publishing On Racial Health Inequities." *Health Affairs Blog*, July 2 2020. <https://www.healthaffairs.org/doi/10.1377/hblog20200630.939347/full/>.
- Brault, Matthew W. 2012. "Americans With Disabilities." Report P70-131. Suitland–Silver Hill, MD: US Bureau of the Census. <https://www.census.gov/library/publications/2012/demo/p70-131.html>.
- Bui, Kim. 2019. "Designing Data Visualisations with Empathy." *DataJournalism.com*, June 11, 2019. <https://datajournalism.com/read/longreads/designing-data-visualisations-with-empathy>.
- Carratala, Sofia, and Connor Maxwell. 2020. "Health Disparities by Race and Ethnicity." Washington, DC: Center for American Progress. <https://www.americanprogress.org/issues/race/reports/2020/05/07/484742/health-disparities-race-ethnicity/>.
- Chartoff, Ben. 2018. "Urban Institute's Data Visualization Style Guide: A Living Document." *Urban Institute Medium* (blog), March 18, 2018. <https://urban-institute.medium.com/urban-institutes-data-visualization-style-guide-a-living-document-b90710702a9f>.
- Cleveland, William S., and Robert McGill. 1984. "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." *Journal of the American Statistical Association* 79 (387): 531–54.
- Costanza-Chock, Sasha. 2020. *Design Justice: Community-Led Practices to Build the Worlds We Need*. Cambridge, MA: MIT Press.

- Correll, Michael. 2019. "Ethical Dimensions of Visualization Research." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, by Association for Computing Machinery. New York: Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300418>
- Del Real, Jose A. 2020. "'Latinx' Hasn't Even Caught On among Latinos. It Never Will." *Washington Post*, December 18, 2020. [https://www.washingtonpost.com/outlook/latinx-latinos-unpopular-gender-term/2020/12/18/bf177c5c-3b41-11eb-9276-ae0ca72729be\\_story.html](https://www.washingtonpost.com/outlook/latinx-latinos-unpopular-gender-term/2020/12/18/bf177c5c-3b41-11eb-9276-ae0ca72729be_story.html)
- D'Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Dörk, Marian, Patrick Feng, Christopher Collins, and Sheelagh Cpendale. 2013. "Critical InfoVis: Exploring the Politics of Visualization." In *CHI 2013: Extended Abstracts on Human Factors in Computing Systems*, edited by Patrick Baudisch, Michel Beaudouin-Lafon, and Wendy E. Mackay, 2189–98. New York: Association for Computing Machinery.
- Dortonne, Nathalie. 2016. "The Dangers of Poverty Porn." *CNN*, December 8 2016. <https://www.cnn.com/2016/12/08/health/poverty-porn-danger-feat>.
- Dougherty, Jack, and Ilya Ilyankou. 2021. *Hands-On Data Visualization*. Newton, MA: O'Reilly Media. <https://handsondataviz.org/>.
- Durkin, Sarah J., Lois Biener, and Melanie A. Wakefield. 2009. "Effects of Different Types of Antismoking Ads on Reducing Disparities in Smoking Cessation among Socioeconomic Subgroups." *American Journal of Public Health* 99 (12): 2217–23.
- Eubanks, Virginia. 2018. *Automating Inequality*. New York: St. Martin's Press.
- Evergreen, Stephanie. n.d. "How Dataviz Can Unintentionally Perpetuate Inequality: The Bleeding Infestation Example." EvergreenData. Accessed February 25, 2021. [http://stephanieevergreen.com/dataviz-inequality\\_pt1/](http://stephanieevergreen.com/dataviz-inequality_pt1/).
- Groeger, Lena V. 2014. "A Big Article about Wee Things." ProPublica, September 25, 2014. <https://www.propublica.org/nerds/a-big-article-about-wee-things>.
- Groeger, Lena V. "When the Designer Shows Up in the Design." ProPublica, April 4, 2017. <https://www.propublica.org/article/when-the-designer-shows-up-in-the-design>
- Groeger, Lena. 2020. "What Coronavirus Job Losses Reveal About Racism in America." ProPublica, July 20, 2020. <https://projects.propublica.org/coronavirus-unemployment/>.
- Hanrahan, Benjamin V., Ning F. Ma, Eber Betanzos, and Saiph Savage. 2020. "Reciprocal Research: Providing Value in Design Research from the Outset in the Rural United States." *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development*, by the Association for Computing Machinery. New York: Association for Computing Machinery.
- Harris, Jacob. 2015. "Connecting with the Dots." Source, January 15, 2015. <https://source.opennews.org/articles/connecting-dots/>.
- Kim, Erika. 2019. "Visible Women: Gender in Iconography." Noun Project blog, October 29, 2019. <https://blog.thenounproject.com/visible-women-gender-in-iconography-c6518ec1ddb8>.
- HEAL Food Alliance. 2020. "An Open Letter from BIPOC Leaders in Food and Agriculture to Food Systems Funders." *HEAL Food Alliance Medium* (blog), July 10, 2020. [https://medium.com/@HEAL\\_Food/an-open-letter-from-bipoc-leaders-in-food-agriculture-to-food-systems-funders-9e00c1f4cb60](https://medium.com/@HEAL_Food/an-open-letter-from-bipoc-leaders-in-food-agriculture-to-food-systems-funders-9e00c1f4cb60).
- Heer, Jeffrey, Michael Bostock, and Vadim Ogievetsky. 2010. "A Tour through the Visualization Zoo." *Communications of the ACM* 53 (6): 59–67.
- Hullman, Jessica. 2019. "Why Authors Don't Visualize Uncertainty." *IEEE Transactions on Visualization and Computer Graphics* 26 (1): 130–39.
- Hullman, Jessica. 2019. "How to Get Better at Embracing Unknowns." *Scientific American*, September 2019. <https://www.scientificamerican.com/article/how-to-get-better-at-embracing-unknowns/>
- Ishisaka, Naomi. 2020. "Why It's Time to Retire the Term 'Asian Pacific Islander'." *Seattle Times*, November 30, 2020. <https://www.seattletimes.com/seattle-news/why-its-time-to-retire-the-term-asian-pacific-islander/>.
- Kaba, Fatos, Angela Solimo, Jasmine Graves, Sarah Glowa-Kollisch, Allison Vise, Ross MacDonald, Anthony Waters, Zachary Rosner, Nathaniel Dickey, Sonia Angell, and Homer Venters. 2015. "Disparities in Mental Health Referral and Diagnosis in the New York City Jail Mental Health Service." *American Journal of Public Health* 105 (9): 1911–16.
- Kendi, Ibram X. 2019. *How to Be an Antiracist*. New York: One World.



- Klein, Scott. n.d. "Intro: The Design and Structure of a News Application." ProPublica GitHub repository, last updated March 4, 2013. <https://github.com/propublica/guides/blob/master/design-structure.md>.
- Lundgard, Alan, Crystal Lee, and Arvind Satyanarayan. 2019. "Sociotechnical Considerations for Accessible Visualization Design." Presentation given at the 2019 IEEE Visualization Conference, October 22–25.
- Lupi, Giorgia. 2017. "Data Humanism: The Revolutionary Future of Data Visualiation." *PrintMag*, January 30, 2017. <https://www.printmag.com/post/data-humanism-future-of-data-visualization>.
- Madden, Mary. 2017. "Privacy, Security, and Digital Inequality." New York: Data & Society Research Institute. [https://datasociety.net/pubs/prv/DataAndSociety\\_PrivacySecurityandDigitalInequality.pdf](https://datasociety.net/pubs/prv/DataAndSociety_PrivacySecurityandDigitalInequality.pdf).
- Marcus, Amy Dockser. 2013. "Trials." *Wall Street Journal*, November 14, 2013. <https://graphics.wsj.com/trials/#chapter=1>.
- Martínez, Alejandro, and Stefano Mammola. 2021. "Specialized Terminology Reduces the number of Citations of Scientific Papers." *Proceedings of the Royal Society B: Biological Sciences* 288 (1948): 20202581.
- MIT (Massachusetts Institute of Technology). n.d. "Diversity Dashboard." MIT Institutional Research, Office of the Provost. Accessed July 2020. <https://ir.mit.edu/diversity-dashboard>.
- McGill, Andrew. 2016. "Why White People Don't Use White Emoji." *Atlantic*, May 9, 2016. <https://www.theatlantic.com/politics/archive/2016/05/white-people-dont-use-white-emoji/481695/>.
- McKernan, Signe-Mary, Caroline Ratcliffe, C. Eugene Steuerle, Caleb Quakenbush, and Emma Kalish. 2015. "Nine Charts about Wealth Inequality (Updated)." Urban Institute, last updated October 5, 2017. <https://apps.urban.org/features/wealth-inequality-charts/>.
- Morais, Luiz, Yvonne Jansen, Nazareno Andrade, and Pierre Dragicevic. 2020. "Showing Data about People: A Design Space of Anthropographics." *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/TVCG.2020.3023013>.
- Myers, Jan, and Martha MacDonald. 2014. "Reciprocal Relationships: The Role of Government and the Social Economy in the Co-Construction of Social Policy in Atlantic Canada." *Canadian Public Policy / Analyse De Politiques* 40 (1S): S17–25. [www.jstor.org/stable/24365086](http://www.jstor.org/stable/24365086).
- Nadjimbadem, Katie. 2017. "The Trashy Beginnings of 'Don't Mess With Texas'." *Smithsonian Magazine*, March 10, 2017. <https://www.smithsonianmag.com/history/trashy-beginnings-dont-mess-texas-180962490/>.
- Nguyen, Francis, X. Qiao, J. Heer, and Jessica Hullman. 2020. "Exploring the Effects of Aggregation Choices on Untrained Visualization Users' Generalizations from Data." *Computer Graphics Forum* 39 (6).
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Noe-Bustamante, Luis, Lauren Mora, and Mark Hugo Lopez. 2020. "About One-in-Four U.S. Hispanics Have Heard of Latinx, but Just 3% Use It." Pew Research Center, August 11, 2020. <https://www.pewresearch.org/hispanic/2020/08/11/about-one-in-four-u-s-hispanics-have-heard-of-latinx-but-just-3-use-it/#fn-29384-5>.
- OMB (Office of Management and Budget). 1977. "Directive No. 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting." Washington, DC: Office of Management and Budget. [https://transition.fcc.gov/Bureaus/OSEC/library/legislative\\_histories/1195.pdf](https://transition.fcc.gov/Bureaus/OSEC/library/legislative_histories/1195.pdf).
- . 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register* 62 (210). <https://www.federalregister.gov/documents/1997/10/30/97-28653/revisions-to-the-standards-for-the-classification-of-federal-data-on-race-and-ethnicity>.
- Onuoha, Mimi. 2016. "The Library of Missing Datasets." MimiOnuoha.com. Accessed May 28, 2021. <https://mimionuoha.com/the-library-of-missing-datasets>.
- Peck, Evan M., Sofia E. Ayuso, and Omar El-Etr. 2019. "Data Is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, by the Association for Computing Machinery. New York: Association for Computing Machinery.

- Porat, Daniela, Lena Groeger, and Isaac Arnsdorf. 2019. "What Happened to All the Jobs Trump Promised?" ProPublica, May 7, 2019. <https://projects.propublica.org/graphics/trump-job-promises>.
- Rodriguez-Lonebear, Desi. 2016. "Building a Data Revolution in Indian Country." In *Indigenous Data Sovereignty: Toward an Agenda*, edited by Tahu Kukutai and John Taylor, 253–72. Canberra: Australian National University, College of Arts and Social Sciences, Centre for Aboriginal Economic Policy Research.
- Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York: Liveright.
- Schwabish, Jonathan. 2021. *Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks*. New York: Columbia University Press.
- Schwabish, Jonathan, and Kilolo Kijakazi. 2021. "Rethinking Economics in the Black Lives Matter Era." Washington, DC: Urban Institute. <https://www.urban.org/research/publication/rethinking-economics-black-lives-matter-era>.
- Shirley, Rachel. 2020. "How Colors Advance and Recede in Art." Science of Color. Accessed July 23, 2020. <https://sites.google.com/site/scienceofcolour/how-colors-advance-and-recede-in-art>.
- Slobin, Sarah. 2014. "What if the Data Visualization Is Actually People?" Source, April 24, 2014. <https://source.opennews.org/articles/what-if-data-visualization-actually-people/>.
- Solomon, Mark G., Richard P. Compton, and David F. Preusser. 2004. "Taking the Click It or Ticket model nationwide." *Journal of Safety Research* 35 (2): 197–201.
- Spriggs, William. 2020. "A Teachable Moment? Will George Floyd's Death Spur Change in Economics?" Federal Reserve Bank of Minneapolis, June 9, 2020. <https://www.minneapolisfed.org/article/2020/a-teachable-moment-will-george-floyds-death-spur-change-in-economics>.
- Squazzoni, Flaminio, Giangiacomo Bravo, Francisco Grimaldo, Daniel Garcia-Costa, Mike Farjam, and Bahar Mehmani. 2020. "Only Second-Class Tickets for Women in the COVID-19 Race. A Study on Manuscript Submissions and Reviews in 2329 Elsevier Journals." Amsterdam: Elsevier SSRN. <https://ssrn.com/abstract=3712813> or <http://dx.doi.org/10.2139/ssrn.3712813>.
- Steenhout, Nicholas. 2020. "Disability Language Is a Nuanced Thing." *Part of a Whole* (blog), July 9, 2020. <https://incl.ca/disability-language-is-a-nuanced-thing/>.
- Tison, Julie, Allan F. Williams, and Preusser Research Group. 2010. *Analyzing the First Years of the Click It or Ticket Mobilizations*. DOT HS 811 232. Washington, DC: US Department of Transportation, National Highway Traffic Safety Administration.
- Tufte, Edward R. 1985. "The Visual Display of Quantitative Information." *Journal for Healthcare Quality* 7 (3): 15. [https://journals.lww.com/jhqonline/citation/1985/07000/the\\_visual\\_display\\_of\\_quantitative\\_information.12.aspx](https://journals.lww.com/jhqonline/citation/1985/07000/the_visual_display_of_quantitative_information.12.aspx)
- Urban Institute. n.d. "Tracking COVID-19's Effects by Race and Ethnicity." Urban Institute. Last updated July 23, 2020. <https://www.urban.org/features/tracking-covid-19s-effects-race-and-ethnicity>.
- US Bureau of the Census. n.d. "About the American Community Survey." US Bureau of the Census. Last revised January 4, 2021. <https://www.census.gov/programs-surveys/acs/about.html>.
- White House. 2021. "Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government." January 20. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>. Wilkerson, Isabel. 2020. *Caste: The Origins of Our Discontents*. New York: Random House.
- Williams, Sarah. 2020. *Data Action: Using Data for Public Good*. Cambridge, MA: MIT Press.
- WebAIM. n.d. "Contrast Checker." WebAIM. Accessed May 12, 2021. <https://webaim.org/resources/contrastchecker/>.

## 19 INTERVIEW GUIDE: DO NO HARM GUIDE AGENDA

---

*Our interviews were conducted between December 2020 and February 2021. Each lasted for about an hour and were conducted via Zoom and transcribed for internal purposes.*

1. Comments or feedback about our existing paper.
  - a. Is there anything we missed? What other topics do we need to discuss and include in the guide?
2. How do you and your organization, or organizations you have worked with, approach taking a racial equity lens to communicating data? This can encompass anything from language to colors to icons to different data visualization types.
  - a. If you've worked on any projects that explicitly applied a racial equity lens to the way you communicated data, we'd love to hear more about that experience. What was the process like? What issues did you consider? What decisions were made? Did you encounter any pushback?
  - b. Can you think of any examples of data projects or data visualizations that did not take a racial equity lens?
3. How do we as data visualizers—and people working with data generally—build relationships and bridges with the people and communities we are visualizing/studying?
  - a. How can we solicit feedback from members of these communities on the design of our visualizations? How should we engage with them as fellow participants in the design process?
    - i. We'd love to hear more about any projects that you have worked on that directly involved the participation of the communities that you were visualizing or studying.
  - b. How can we sustain these relationships so that they last beyond the duration of any single project?
4. How do internal organizational culture and analysis overlap?
  - a. How can data visualization and data science teams (and the larger organizations they're housed in) do a better job of reflecting the diverse experiences of the communities we are visualizing/studying?
  - b. Are there aspects of the process of designing and creating a data visualization that could be improved on to be more diverse, equitable, and inclusive?
5. Is there anything else you'd like to add? Anything we missed? Do you have any recommendations for other people we should speak with?

## 20 OTHER DATA AND DATA VISUALIZATION RACIAL EQUITY RESOURCES

Below we have listed a sampling of books, organizations, and other resources (in addition to those found in the references section) that may be useful to help you and your organization build an equitable approach to data and data visualization. This list will be published on the Racial Equity Data Hub from the Tableau Foundation.

### Selected Bibliography

- Catherine D'Ignazio and Lauren F. Klein, *Data Feminism* (Cambridge, MA: MIT Press, 2020).
- Sasha Costanza-Chock, *Design Justice: Community-Led Practices to Build the Worlds We Need* (Cambridge, MA: MIT Press, 2020).
- Sarah Williams, *Data Action: Using Data for Public Good* (Cambridge, MA: MIT Press, 2020).
- G. Cristina Mora, *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American* (Chicago: University of Chicago Press, 2014).
- Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: NYU Press, 2018).
- Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code* (Cambridge, UK: Polity Press, 2019).
- Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin's Press, 2018).
- Witney Battle-Baptiste and Britt Rusert, eds., *WEB Du Bois's Data Portraits: Visualizing Black America* (Hudson, NY: Princeton Architectural Press, 2018).
- Tactical Technology Collective, *Visualizing Information for Advocacy* (Berlin: Tactical Technology Collective, 2013), <https://visualisingadvocacy.org/get.html>.

### Organizations

- Allied Media Projects, <https://alliedmedia.org/>
- Design Justice Network, <https://designjustice.org/>
- Detroit Digital Justice Coalition, <http://detroitdjcc.org/>
- The A11Y Project, <https://www.a11yproject.com/>
- We All Count, <https://weallcount.com/>
- Civic Data Lab (MIT), <https://civicdatadesignlab.mit.edu/>

- The Communications Network, <https://comnetworkdei.org/>
- Black Futures Lab, <https://blackfutureslab.org/>
- Data for Black Lives, <https://d4bl.org/>
- The Marshall Project, <https://www.themarshallproject.org/>
- Teachers 4 Social Justice, <https://t4sj.org/>
- Chartability, <https://chartability.fizz.studio/>

### Other Resources

- Child Trends' *How to Embed a Racial and Ethnic Equity Perspective in Research: Practical Guidance for the Research Process*, [https://www.childtrends.org/wp-content/uploads/2019/09/RacialEthnicEquityPerspective\\_ChildTrends\\_October2019.pdf](https://www.childtrends.org/wp-content/uploads/2019/09/RacialEthnicEquityPerspective_ChildTrends_October2019.pdf)
- Actionable Intelligence for Social Policy's *A Toolkit for Centering Racial Equity Throughout Data Integration*, <https://www.aecf.org/resources/a-toolkit-for-centering-racial-equity-within-data-integration/>
- Chicago Beyond's *Why Am I Always Being Researched*, <https://chicagobeyond.org/researchequity/>
- Black Design in America, <https://bipocdesignhistory.com/>
- *Scratching the Surface* podcast, <https://scratchingthesurface.fm/>
- Urban Institute
  - "Taking an Equity Lens to our Data Practice," <https://medium.com/local-data-for-equitable-recovery/taking-an-equity-lens-to-our-data-practice-8a980fe14d0>
  - "Principles for Advancing Equitable Data Practices," <https://www.urban.org/research/publication/principles-advancing-equitable-data-practice>
  - "Equitable Data Practice," <https://www.urban.org/elevate-data-equity/resources-elevate-data-equity/equitable-data-practice>

## 21 DIVERSITY, EQUITY, AND INCLUSION IN DATA VISUALIZATION: GENERAL RECOMMENDATIONS

---

- **Critically examine your data.**
  - Understand where your data come from, who is included and excluded from these data, how these data were collected, why they were collected, and who benefits or is harmed by them.
- **Use people-first language.**
  - Start with the person, not the characteristic such as “people with disabilities,” “a person with asthma,” or “communities of color.”
- **Label people, not skin color.**
  - Use full labels such as “Black people” rather than “Black.”
  - Remember that language continues to evolve. Certain labels that may have been acceptable years ago may no longer be.
- **Order labels purposefully.**
  - Don’t simply order data in tables, graphs, and charts as they are ordered in the data, which may reflect historical biases. Consider alternative sorting parameters such as study focus, specific story or argument, quantitative relationship (i.e., magnitude of the results), alphabetical, or sample size (weighted or unweighted).
- **Consider missing groups.**
  - What groups are not included in your data? Consider adding notes to highlight how the data are not inclusive or representative.
  - Consider alternatives to labeling the “other” catch-all category:
    - *Another*
    - *Another race*
    - *Additional groups*
    - *All other self-descriptions*
    - *People identifying as other or multiple races*
    - *Identity not listed*
    - *Identity not listed in the survey*
- **Carefully consider colors.**
  - Avoid reinforcing gender or racial stereotypes, such as by using baby pink and baby blue to represent women and men or colors associated with skin tones or racial stereotypes.
  - Avoid using incremental color palettes (e.g., light to dark) to represent different demographic groups.
- **Consider icons and shapes.**
  - Recognize how readers might be better able to connect with the data by using small shapes or icons, but use them carefully.
  - Avoid stereotypical, discriminatory, and racist imagery.
  - Use images that show people as empowered and dignified, and avoid images that depict people as helpless victims.
- **Communicate with people and communities of focus.**
  - Reach out to the people and communities you are focusing on in your work and hoping to connect with through the final product.
  - Build teams and connections with outside groups to build and maintain these relationships. This kind of work takes time and effort.
- **Reflect lived experiences.**
  - Not everyone has the same experiences, especially when it comes to characteristics such as race, ethnicity, and gender. Consider what your work may be missing and seek out colleagues and communities to help identify them.
- **Consider the needs of your audiences.**
  - Make sure results are presented in a format that is useful to the audience.
  - Make sure the language used is written in a way that is easily understandable by your readers.
  - Consider translating your products into the languages used by your audiences.

## 22 THE RACIAL EQUITY IN DATA VISUALIZATION CHECKLIST

---

- |   |   |
|---|---|
| <ul style="list-style-type: none"> <li><input type="checkbox"/> Does the communicator understand the data they worked with, including how it was sourced, who was or was not represented in it, why it was collected, and who benefits or is harmed by having these data collected?</li> <li><input type="checkbox"/> Has the communicator carefully considered words, phrases, and labels that are used to describe people, groups, and communities?</li> <li><input type="checkbox"/> Has the communicator considered colors that are inclusive of different groups and that are accessible for people with different abilities?</li> <li><input type="checkbox"/> Has the communicator considered the order of numbers or estimates in tables, charts, and diagrams? Some options include sorting alphabetically or by magnitude of estimate/number, population size, and sample size (weighted or unweighted).</li> <li><input type="checkbox"/> Has the communicator considered alternative words or phrases for groups that may be classified as “other” in the original data? Some alternatives include another race; additional groups; all other self-descriptions; people identifying as other or multiple races; identity not listed; and identity not listed in the survey.</li> <li><input type="checkbox"/> Have all icons and images been reviewed with a racial equity lens?</li> <li><input type="checkbox"/> Would alternative graph types do a better job presenting the data? Do all groups need to be positioned within the same graph?</li> </ul> | <ul style="list-style-type: none"> <li><input type="checkbox"/> Has the research team communicated with the people or communities that they are focusing on or wishing to communicate with? If not, what people, groups, or organizations can the team contact?</li> <li><input type="checkbox"/> Is the research team—and the organization more generally—diverse in its composition and work practices to be able to facilitate better understanding of different groups?</li> <li><input type="checkbox"/> Does the final communication product meet the needs of the audience or user?</li> </ul> |
|---|---|

**Teams should consider these issues throughout the research and communication process. Use these checkboxes as a reminder:**

- ☐ *Proposal development stage*
- ☐ *Strategic planning stage*
- ☐ *Data collection phase*
- ☐ *Analysis phase*
- ☐ *Data visualization phase*
- ☐ *Writing phase*
- ☐ *Editing phase*
- ☐ *Outreach and final publication stage*



## ABOUT THE PROJECT

This guide was funded by the Tableau Foundation. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission. The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at [urban.org/fundingprinciples](https://urban.org/fundingprinciples).

## ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.

Copyright © June 2021. Urban Institute.

Permission is granted for reproduction of this file, with attribution to the Urban Institute.