Kevin Donnelly
Extended Abstract: Workshop on Biases in Social Computing Data and Technology

Deviance and Bias in Nineteenth-Century Social Physics and Anthropometry

This paper aims to provide a historical context in which to appreciate the shared problems faced by social and computer scientists in combating bias. While historians of science are well aware of the trials (and errors) of previous attempts to quantify the human condition, this literature has not always made it into discussions of modern CSS. Indeed, some manuals for the computational social sciences often imagine that data mining and statistical averages are new, or that "Big Data" has only existed for the past decade. Popular books on the subject rarely mention that their ideas are far from novel, and that claims for a "quantitative moment" are centuries old. Yet historians have noted that the "avalanche" of numbers occurred close to 200 years ago, and that a "probabilistic revolution" radically altered how people imagined human actions. Such selective historical amnesia has led, this paper argues, to problems in combatting bias today. Seen only in the light of present politics and practical concerns, these problems appear intractable, or technical, yet the historical record reveals numerous instances of bias emerging, and being addressed, in the quantitative social sciences.

To help provide historical context to the idea of bias in numbers, the paper will examine controversies that arose over two of the earliest attempts at a computational social science: the dual projects of social physics and anthropometry created by Adolphe Quetelet, the "father of modern statistics." While attempting to construct a science of society in the early nineteenth century, Quetelet relied heavily on deterministic Laplacian metaphors, believing that what he called his "average man" would become an abstract construct that would serve as a center of gravity around which all human behavior revolved. So strong was the analogy that Quetelet called his project *physique sociale* or *mécanique sociale* in order to stress its predictability. Deviations in human social behavior were then thought of as mere deviations of measurement and design, and Quetelet imagined only the averages as consistent with the wishes of God and nature. While he explicitly modeled his ideas on the French Enlightenment's vision of a harmonious society, it was nevertheless appropriated by later British and American social scientists more interested in Darwin's competition than Laplace's stability. By grafting new biological metaphors onto a concept rooted in enlightened visions of science, these social scientists helped to create a hybrid of social and biological sciences that transformed the very concept of average into a measure of deviance.

Yet such a focus on deviance and anomaly was not the goal of nineteenth-century social physics. While modern CSS has moved far beyond Quetelet's crude averages, the idea of designing a data set and reporting protocol in order to eliminate bias has remained quite similar. Quetelet is certainly known in academic circles for his abstract ideas and legacy in the quantitative social sciences, but this paper will investigate how he attempted to create the practical institutions (and workers) to carry out his dream of collecting data on all aspects of human behavior. For those actions he could not record, he proposed a myriad of observational techniques and analog algorithms, including forms with specific steps for his workers to follow in creating what today we might call "data doubles." In a peculiar symmetry, Quetelet sought out real average men to do the work of surveillance on his abstract "average man," screening out workers that deviated

from the norm. In stressing the connections between data collection and data collectors, this paper examines the late-nineteenth century complaints about the tyranny of numerical surveillance. In doing so, it hopes to offer lessons for understanding our own issues with quantitative observation.

In order to open up the conceptual space to explore the connections in the history of the statistical social sciences and modern CSS, the paper draws on a relatively new field of study known as the sociology of quantification. Building on the groundbreaking work of Alain Desrosières, Pierre Bourdieu, and Bruno Latour, this field examines what the process of counting and quantifying means to the actors involved in the process as well as those being counted. Is there a particular set of interactions that must occur in order for certain things to "count"? What is the tacit knowledge of the quantifiers? Can there be such a thing as "raw data"? Or, are the preoccupations and interests of those doing the counting "built in" to the numbers themselves? Through an investigation of these questions and the praxis of counting, it helps to explain how the quantifiers from the Enlightenment up until today often share similar concerns and dispositions towards their numbers and subjects. By showing how bias had been combatted and addressed in the past, it hopes to offer possibilities for modern researchers vexed by the possibility of implicit error in their data.