MAKING AND UNMAKING POPULATIONS[*]

Staffan Müller-Wille

Egenis, the Centre for the Study of the Life Sciences

University of Exeter

Byrne House

St Germans Road

Exeter, Devon EX4 4PJ

United Kingdom

sewm201@ex.ac.uk

[*] Accepted for publication in a special issue on "Histories of Data and the Data Base" guest-edited by Soraya de Chadarevian and Ted Porter for *Historical Studies of the Natural Sciences* 48, no. 5 (2018).

ABSTRACT

Statistics derives its power from classifying data and comparing the resulting distributions. In this paper, I will use two historical examples to highlight the importance of such data practices for statistical reasoning. The two examples I will explore are Franz Boas's anthropometric studies of native American populations in the early 1890s which laid the foundations for his later critique of the race concept; and Wilhelm Johannsen's experiments in barley breeding, which he carried out for the Carlsberg Laboratory around the same time and which prepared the ground for his later distinction of genotype and phenotype. Both examples will show that the manipulation of data depended on complex classificatory practices: the distinction and articulation of "tribes," "races" and "family lines" in the case of Boas, and the selection and construction of "populations" and "pure lines" in the case of Johannsen. They also reveal a fundamental difference between data practices in the human and the life sciences: Whereas the latter are relatively free to construct populations in the laboratory, the field, or on paper, the former have to rely on social categories shaped by historical accident and self-perception of the subjects under study.

KEYWORDS: Franz Boas (1858–1942); Wilhelm Johannsen (1857–1927); history of statistics; history of genetics; racial anthropology; anthropometry; plant breeding; Indian Territories

[FIRST LEVEL HEADING] INTRODUCTION

In the twentieth century, “statistics” has come to mean a branch of applied mathematics dealing with the analysis of large amounts of numerical data. According to Ted Porter’s classic history of the discipline, this sense of statistics emerged in the 1830s and 1840s in association with the “great explosion of numbers” caused by the bureaucratization of European nation states.[2] Some early definitions of statistics, however, draw attention to an aspect in addition to large numbers. Thus political economist Francis Ysidro Edgeworth (1845–1926) declared in 1885 that statistics was a “science of Means,” where “‘Means’ imply the correlative conception of terms of a ‘series,’ or members of a class, whose mean is to be taken.”[3]

Statistics was thus not simply concerned with numbers, but with numbers relating to things classified. Statistical pioneers like Adolphe Quetelet (1796–1874), Francis Galton (1822–1911), Wilhelm Lexis (1837–1914) or Karl Pearson (1857–1936) were not interested in populations composed of homogenous individuals exhibiting well-defined differences like the colored balls in urn models. They were interested in (mostly human) populations made up from individuals that differed by geographic location, age, family position, gender,

---

[2] Theodore M Porter, *The Rise of Statistical Thinking: 1820-1900* (Princeton: Princeton University Press, 1986), 3, 11.

[3] Syllabus for Edgeworth’s 1885 Lectures, King’s College, London, quoted in Stephen M. Stigler, *The History of Statistics: The Measurement of Uncertainty Before 1900* (Cambridge, MA: Harvard University Press, 1986), 363.

profession, class or race.[4] The data they manipulated were therefore always already structured by categories that reflected the ontologies of census-taking.[5] While Edgeworth showed his penchant for classical knowledge when he chose to pursue his analysis of "explixit" and "entangled fluctuations"—an early precursor of analysis of variance—on the basis of counting dactyls per foot and line of a section from Virgil's *Aeneas*, the structural analogy between this example and census data organized by place and year was not far-fetched.[6]

Statistics thus involves more than mathematics. To make this point, I will focus on two figures that are hardly ever mentioned in the historiography of statistics, but whose statistical work had a huge impact on their respective disciplines. In a first section, I will look at German-American anthropologist Franz Boas's (1858–1942) anthropometric studies of Native American tribes in the early 1890s. These studies provided the groundwork for his sustained critique of the race concept.[7] The second section will focus on breeding experiments that the Danish botanist Wilhelm Johannsen (1857–1927) carried out in the mid-1890s and that foreshadowed his later attack on traditional understandings of biological

---

[4] Theodore M. Porter, „Statistics and Statistical Methods," in *The Modern Social Sciences*, ed. Theodore M. Porter and Dorothy Ross, (Cambridge: Cambridge University Press, 2003), 238–50.

[5] On census taking and the history of statistics, see Margo J. Anderson, *The American Census: a Social History* (New Haven, CT: Yale University Press, 1988); Eric Brian, *La Mesure de l'état: administrateurs et géomètres au XVIIIe siècle* (Paris: Albin Michel, 1994).

[6] Stephen M. Stigler, "Francis Ysidro Edgeworth, Statistician," *Journal of the Royal Statistical Society* A141 (1978): 299–304.

[7] George W. Stocking, "The Critique of Racial Formalism," in *Race, Culture, and Evolution: Essays in the History of Anthropology* (New York: The Free Press, 1968), 161–94.

inheritance.[8] In both cases, as I will show, it was the skillful deployment of data practices to make, and unmake, populations, rather than mathematical proficiency, that allowed Boas and Johannsen to take their respective disciplines a critical step further.

[FIRST LEVEL HEADING] BOAS: TRIBES, RACES AND FAMILY LINES

In 1892, Frederic Ward Putnam (1839–1915) hired Boas to assist him in preparing an anthropology exhibition at the World's Columbian Exposition to be held in Chicago in the following year. One of the fields on which Boas focused was the physical anthropology of Native Americans, or "Indians", as he called them. In previous years, he had already conducted small-scale anthropometric campaigns, and now he intensified this activity. With about 50 field observers, he collected data on c. 10,000 individuals, focusing on "tribes" inhabiting the East Coast and the Indian Territories.[9]

A large number of original data sheets from these surveys have been preserved. The data sheets are remarkable, above all, for a first section that included detailed genealogical information on each person, and which was probably inspired by the forms used in the 1890 US Census. Not only name, location and age were recorded, but also tribal affiliation, tribal affiliation of father and

---

[8] Staffan Müller-Wille and Marsha L. Richmond, "Revisiting the Origin of Genetics," in *Heredity Explored: Between Public Domain and Experimental Science, 1850-1930*, ed. Staffan Müller-Wille and Christina Brandt (Cambridge, MA: MIT Press, 2016), 375–81.

[9] Richard L. Jantz, "Franz Boas and Native American Biological Variability," *Human Biology* 67 (1995): 345–53.

mother, as well as family relationship to other individuals covered by the survey. The details filled in for "Tribe of mother" and "Tribe of father" respectively often refer to complex ancestral backgrounds; on many of the sheets I analyzed, one finds information like "half-breed", "Tribe X and Negro," or "¾ Tribe X ¼ white" for "tribe of mother" or "father," respectively. This information reflected what the subjects of the survey said about themselves, rather than being based on judgments made by the field assistants or Boas himself.

The two sections that followed on the data sheets turned to anthropometric variables. Observations on qualitative traits like eye or hair color and a series of twelve anthropometric measurements restricted to stature and head form were filled in by the field assistants who had been trained by Boas to eliminate inter-observer variability as far as possible. This section was accordingly filled out in a mechanical manner, leaving no room for the assistant to add any observations of their own. A final and third section was reserved for Boas to calculate "indices," such as the cephalic index, or the ratio between breadth and length of head. In this, he proceeded highly selectively; he did not calculate each index for each individual, but particular indices for particular subsets of individuals only, such as "full-blooded" members of tribes.[10]

Boas summarized results from the anthropometric surveys in a long article, which appeared in 1895 and made liberal use of tables and curve diagrams. He

[10] For a detailed reconstruction of how Boas's data were collected, see Staffan Müller-Wille, "Joining Data Across Cultures: Kinship Analysis and Statistics in Late Nineteenth-Century Anthropology," in *Varieties of Data Journeys: Data Processing and Movements Within and Across Practices*, ed. Sabina Leonelli and Niccolo Tempini (Cambridge, MA: MIT Press, in preparation).

admitted right away that data on qualitative traits varied too much depending on observers to deliver comparable results, and for this reason focused exclusively on three quantitative traits, namely body height, cephalic index and facial width.[11] This poverty of variables was compensated, however, by the virtuosity with which Boas considered the distribution of variables in relation to geography, age, gender, tribal affiliation and "mixed" ancestry. For example, by comparing individuals of the same tribal affiliation, but living in different geographic areas—many of the "tribes" had been forcibly removed—Boas sought to demonstrate environmental influences, while individuals of different tribal affiliation, yet living in the same region, provided evidence for "hereditary influences."[12] A fact that fascinated him in particular was that distribution curves of "half-blooded" populations did not show simple blending of the two parental types, but usually two maxima, and hence "regression" to each of the two parental types.[13] He even tried to demonstrate—by "classifying mixed-bloods in such a way, that one group includes individuals which have more than half of Indian blood, and the other individuals, which have half or less Indian blood"—that the "Indian type" possessed a "stronger hereditary force" (*grössere Vererbungskraft*).[14]

[11] Franz Boas, "Zur Anthropologie der nordamerikanischen Indianer," *Zeitschrift für Ethnologie* 27 (1895): 367.

[12] Ibid.: 375–376.

[13] Ibid.: 406–7.

[14] Ibid.: 410–11.

A striking example of Boas's versality in constructing and deconstructing populations on paper can be found in an article that he published under the ambitious title "The Correlation of Anatomical or Physiological Measurements." After a long-winded mathematical discussion of correlation, Boas abruptly moved on to provide "an illustration." Selecting data for "length and breadth of head of 923 adult male Sioux, Crow, and western Ojibwa," Boas constructed two tables, one breaking down the pooled population into subpopulations classed by length of head in steps of five millimeters, the other doing the same by breadth of head, and then calculated the averages of both measures for each of the resulting subpopulations.[15] In addition, he plotted these averages in a curve diagram, resulting in two more or less straight lines that formed a conspicuous cross (see Figure 1). That a particular length of head has no appreciable effect on breadth of head, and vice versa, and that the two measures hence vary, is immediately obvious from this curve diagram. The cephalic index, as Boas suggested against a widely held assumption, could therefore "not indicate any biological law."[16]

[PLACE FIGURE 1 HERE]

After 1901, Boas abandoned his anthropometric surveys among Native Americans. However, he followed a similar data collection design in the famous large-scale anthropometric study he carried out for the U.S. Immigration Commission in 1910, which showed that physical characteristics changed slightly but significantly in American-born children of immigrants of different

[15] Franz Boas, "The Correlation of Anatomical or Physiological Measurements," *American Anthropologist* 7 (1894): 317.

[16] Ibid.: 314.

ethnic origin or "race" (Eastern European Jews, Sicilians, Neapolitans, Czechs, Poles and Scots). Boas and his assistants collected data from almost 18,000 individuals using a form that was similar to the one discussed above and hence also allowed to compare data for parents and their children.[17] As he later emphasized, there was a tension between such genealogical units— "genetic" or "family lines," as he called them—and the concept of race. The "races" of everyday language usually did not refer to homogenous "types", but to aggregates of family lines, and vice versa, family lines were often composed of individuals belonging to different racial types.[18] A proper, statistical study of biological aspects of human populations therefore had to break them down into their constituent family lines, not into racial types, as most physical anthropologists believed at the time.[19]

[FIRST LEVEL HEADING] JOHANNSEN: POPULATIONS AND PURE LINES

In the same year that the World's Columbian Exposition took place in Chicago, the Danish botanist Wilhelm Johannsen, then teaching at the Royal Agricultural and Veterinary College in Copenhagen, started a series of breeding experiments. Prior to his appointment at the agricultural college, he had been working in the biochemistry department of the Carlsberg Laboratories, carrying out

[17] Franz Boas, *Changes in Bodily Form of Descendants of Immigrants. (Final Report)* (Washington: Government Printing Office, 1911), 81, 117–28.

[18] Franz Boas, "On the Variety of Lines of Descent Represented in a Population," *American Anthropologist* 18 (1916): 2; Franz Boas, "Report on an Anthropometric Investigation of the Population of the United States," *Journal of the American Statistical Association* 18 (1922): 197.

[19] Franz Boas, *The Mind of Primitive Man*, revised edition (New York: Macmillan, 1938), 55.

experiments on the role of organic nitrogen in the maturation of seed. The trial plots and workforce necessary to carry out breeding experiments were provided by Knuthenborg Avlsgaard, a large estate near Copenhagen, which had been used for wheat and barley breeding experiments by the Danish Agricultural Association since 1889.

Johannsen's experiments focused on two economically important properties of barley seeds: their weight and their nitrogen content, both traits of commercial significance. He began by selecting eighty-six spikes of the commercial variety "Carter's Goldthorpe." Some seeds from the spikes were used as samples to determine seed weight and nitrogen content, while the others were sown out to produce the next generation. This was repeated in the following year, but in the third year of the experiment, Johannsen switched to breaking down the population into genealogically defined subpopulations—"lines" (*rækker*), as he called them, or pedigrees, as we would call them today. In this he was following a method that the French plant breeder Louis de Vilmorin (1816–1860) had developed some 50 years earlier. The method consisted in raising all descendants from one and the same individual seed on separate trial plots while taking measures to prevent any cross-fertilization or admixture.[20] This only

[20] Wilhelm Johannsen, "Fortsatte studier over kornsorterne I. Om variabiliteten med særligt hensyn til forholdet mellem kornvægt og kvæfstof-procent i byg," *Meddelelser fra Carlsberg Laboratoriet* 4 (1899): 244.

worked for plants that were self-fertilizing like barley and hence, as Johannsen liked to put it, never left a doubt about their "father."[21]

After carrying through these experiments to the fourth generation—the generation of "grandchildren" (*børnebørn*) of the individual seeds selected for pedigree breeding in the second generation[22]—it took Johannsen another three years of processing the data before he could present his results in the journal that was published by the Carlsberg Laboratory. For each generation, a "table of details" recorded seed weight and nitrogen for plant individuals grown from seed harvested in the previous year. Alongside this, he put together tables in which the degree of covariation between seed weight and nitrogen content was made evident by classifying the material according to these variables, and then establishing averages and absolute frequencies against this classification. In the first year (1893), with its random selection of spikes, these tables indicated quite a strong positive correlation between the two variables. In the second year (1894), which was based on a selection of spikes combining high weight with low nitrogen content, and vice versa, correlation was even stronger and variation less pronounced. In the third year, Johannsen switched to pedigree breeding. Selecting individual seeds from twenty-five spikes that had represented "outliers" (*Undtagelser*) combining high weight with low nitrogen in the previous generation, he raised plants from these seeds on carefully isolated plots.[23] The

[21] Nils Roll-Hansen, "Sources of Wilhelm Johannsen's Genotype Theory," *Journal of the History of Biology* 42, (2009): 476.

[22] Johannsen, "Fortsatte Studier" (ref. 19): 263.

[23] Ibid.: 244.

data for the seed harvested from these plants revealed some considerable differences between the "lines"; not only did positive correlation vary in its strength, there even was one case, in which the two variables seemed to be negatively correlated. Johannsen summarized this result in a diagram that visualized covariation in the form of variously inclined lines that plotted average weight against average nitrogen content for five weight classes.

This result was of considerable practical and theoretical significance. According to a view widespread among both breeders and biologists, the intricate physiological interdependencies in an organism led to "laws of correlation" between certain traits. And one such law claimed that high seed weight in barley always went along with high nitrogen content. This was unfortunate, since low nitrogen content had a positive effect on the malting quality of barley.[24] What Johannsen could show was that the presumed law was an aggregate property of populations that, if appropriately "unmade", i.e. divided into their constituent "lines," could be shown to harbor other laws. In order to prove this point beyond doubt, Johannsen pooled all the data from the descendants of the 25 "outliers" so carefully separated in 1895 and produced a graph that indicated *no* correlation at all (see Figure 2). "The image reminds one of a star map," Johannsen explained. "But if the 'law' of a rise in nitrogen percentage with grain weight was strictly typical, one would have to find a dense 'Milky Way' of points ascending from left to right."[25]

---

[24] Nils Roll-Hansen, "Sources" (ref. 20): 473–74.

[25] Johannsen, "Fortsatte Studier" (ref. 19): 262.

[PLACE FIGURE 2 HERE]

Like Boas, Johannsen was thus lead to believe that populations of living beings formed complex aggregates that needed to be analyzed genealogically. This was the general point of a short book entitled “Heritability in Populations and Pure Lines” (*Ueber Erblichkeit in Populationen und in reinen Linien*) that Johannsen published in 1903. The subject of the book was the “law of regression” that Galton had proposed in 1885 and that had become one of the central tenets of the Biometric school led by Pearson. According to this law, selection in a parental generation would always lead to a partial regression of offspring towards the average type of the overall population. On the basis of experiments with Princess Beans, in which he employed the same techniques of decomposing and recomposing populations as in his earlier barley studies, Johannsen demonstrated that the law of regression was only valid for genetically heterogeneous populations. In “pure lines” produced by pedigree breeding, in contrast, regression was not partial but “complete, all the way to the type of the line.”[26] Johannsen concluded from this that the “personal constitution of parents, grandparents or any ancestor ... has no influence on the average character of descendants;” instead, it was the “type of the line that determines the average character of descendants.”[27] This conclusion anticipated his later distinction of phenotype and genotype, and constituted a direct attack on traditional notions of inheritance.

[26] Wilhelm Johannsen, *Ueber Erblichkeit in Populationen und in reinen Linien* (Jena: Gustaf Fischer, 1903), 39.

[27] Ibid., 61–62.

[FIRST LEVEL HEADING] CONCLUSION

The two case studies presented in the preceding suggest two historiographical conclusions. The first concerns the epistemic role of classification in data-intensive sciences. The examples of Boas and Johannsen show that this role is not restricted to the theoretical role of supplying categories or "lables" by which data are interpreted, disseminated and integrated, as argued by Sabina Leonelli.[28] In addition, classification plays an empirical role when scientists collect, manipulate, and process data in order to present empirical findings in forms, tables and graphs. Filling out a form, drawing up a table or plotting a graph are activities that in themselves involve the grouping of elements in particular ways, and hence classification.[29] In other words, there is no data without a "data-base," i.e. without an infrastructure that contains and classifies data in some way, yet, crucially, also allows for their extraction and re-classification.[30]

This means that surveys and experiments like those discussed in this paper will depend on prior ontological assumptions about the structure of populations, and that these assumptions can often be culturally entrenched. This is particularly evident in the case of Boas. In stark contrast to Johannsen, he was not able to

[28] Sabina Leonelli, *Data-Centric Biology: A Philosophical Study* (Chicago: University of Chicago Press, 2016), ch. 5, esp. 128.

[29] On the empirical primacy of classification, see John Dupré, "In Defence of Classification," *Studies in History and Philosophy of Biological and Biomedical Sciences* 32 (2001): 203–19.

[30] Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge: The MIT Press, 1999).

control the reproduction of the populations he observed. He had to rely on information about tribal affiliation, racial ancestry, and family position that the subjects of his surveys provided in response to corresponding questions and that were shaped by political contingencies. This did not prevent him, however, from making novel empirical claims about the social categories, or "interactive kinds," he was operating with, by breaking up and pooling his data in a variety of ways. In a sense, of course, Johannsen's barley plants "responded" as well to their treatment, but by breaking down populations into isolated "lines," he was able to construct population structures with far greater liberty in order to uncover hidden empirical regularities.[31]

The second conclusion comes back to the historiography of statistics with which I opened this essay. It is often noted that statistical parameters like "regression," "correlation," or "heritability" originally had a biological meaning, and that this shows how intimately the history of statistics was entwined with the history of the sciences in which it was applied.[32] Innovation in the history of statistics is therefore not only associated with the introduction of new statistical parameters, as important as they are. Equally important were innovations with regard to the design of surveys or experiments and with regard to the way in which data were selected from the resultant repositories in order to present them in tables and

---

[31] On Ian Hacking's notion of interactive kinds, see Muhammad Ali Khalidi, "Interactive Kinds," *The British Journal for the Philosophy of Science* 61 (2010): 335–60, who argues that they are not restricted to the human sciences.

[32] See Staffan Müller-Wille and Hans-Jörg Rheinberger, *A Cultural History of Heredity* (Chicago, 2012), 109–110.

graphs. Most importantly, such data practices always leave scope for "making and unmaking" populations in unprecedented and surprising ways, as we saw in the case of Boas and Johannsen.

In order to understand statistical reasoning and its history, we therefore need to attend to data practices, and we need to do so, not only to understand the role of statistics in the "verification," or else "falsification," of statements conforming to already existing ontologies, but also its critical role in the "rectification" of such ontologies, i.e. their transformation by entirely new ontologies.[33] Both Boas and Johannsen, it should be noted, carried out the statistical studies I analyzed above almost two decades before the concept of the "gene" was introduced. And yet, with the benefit of hindsight, one can see how their studies anticipated this new entity. That statistics involves more than mathematics works both ways; it presupposes more than mathematics, and it produces more than mathematics.

[FIRST LEVEL HEADING] ACKNOWLEDGEMENTS

I would like to thank the organizers and participants of the conference "Histories of Data and the Database" at The Huntington Library, San Marino, CA, November 18 and 19, 2016, for the opportunity to present and discuss an early version of this paper. Special thanks go to Sara Green and Claus Emmeche, who invited me in October 2017 to Copenhagen to discuss this paper in the science studies series at Niels Bohr Institute, and to Lene Koch who posed some trenchant questions

[33] On Gaston Bachelard's notion of "rectification," see Hans-Jörg Rheinberger, "Gaston Bachelard and the Notion of 'Phenomenotechnique,'" *Perspectives on Science* 13 (2005): 313–28.

on this occasion. Elena Isayev provided some very precious advice on footnoting, grammar and Virgilian hexameters.