7
Shaping the Unruly Statistician

Theodore M. Porter

In his *Grammar of Science*, first published in 1892, Karl Pearson emphasized that accurate classification of facts was the first step of scientific method.[1] He was just beginning at this point to think of himself as a statistician, and of scientific method as closely linked to statistics. A case could be made for Pearson as the first modern statistician, yet his field has always been heterogeneous, even ill-defined, resisting any neat definition. What Pearson may have founded was a mathematical field. *Statistics* had already been around under other definitions for more than a century. He could be quite critical of these predecessor forms of statistics, yet he did not want to sacrifice any of their breadth.

Even if we ignore the etymology of this *state-istics* and insist on a definition in terms of quantified knowledge, the *statistician* remains an elusive quarry. Only since the 1930s has it been reasonably possible to take an advanced degree in this kind of statistics. Meanwhile, many practitioners have focused their study on one or more substantive disciplines including social science, astronomy, economy, demography, natural history, psychology, evolutionary biology, and eugenics, where much statistical teaching also has taken place. The work of official statistics, most notably the census, cannot be neatly excised from this field, since it draws heavily on mathematical tools of data preparation. Much of the work of statistics consists of service to other disciplines such as medicine, engineering, and business. There can be no neat history of the professional training of statisticians except one that is oversimplified to the point of falsehood. This essay addresses statistics as the foundation for a broad ecology of enumeration, inference, and measurement.

This paper departs in several respects from the classic story of the nineteenth-century German university, which first made science and scholarship into the basis for a career. Pearson in 1892 remained a

[1] Karl Pearson, *The Grammar of Science* (London, 1892), 8.

Theodore M. Porter, *Shaping the Unruly Statistician* In: *A Global History of Research Education: Disciplines, Institutions, and Nations, 1840–1950*. Edited by: Ku-ming (Kevin) Chang and Alan Rocke. *History of Universities* XXXIV/1, Oxford University Press (2021).
DOI: 10.1093/oso/9780192844774.003.0008

fervent advocate of this German ideal, which treated the inculcation of research ideals through close study of a particular subject as a form of self-cultivation. In practice, it ordinarily led to disciplinary specialization. While the Germans were not alone in sharpening their focus on discipline-based knowledge, the new research university provided an environment for a more systematic specialist training. Yet the university ideal could never have survived if it had been simply inward-looking. Quite apart from their general role as a marker of class and culture, advanced degrees were linked to a variety of careers. Studies involving ancient languages or mathematics, for example, were integral to the preparation of *Gymnasium* (secondary school) teachers. In the natural sciences, and especially in technological fields, training at universities led also to research positions in private industry. Later in the century, university research institutes in fields like chemistry were often devoted to industry and technology more than to education. The higher faculties of law, medicine, and theology trained students for professional roles and generally shunned the focus or narrowing required by a research specialty. Research and training in practical fields, including engineering, mining, and agriculture, resembled the professions in being organized to serve clients or businesses of various sorts, and only secondarily to cultivate new knowledge in a discipline.[2]

Statistics, whether as a substantive or as a methodological study, is especially difficult to pin down. It was originally understood as an empirical science of the state, and in that form it was practiced by scholars and state officials beginning in the eighteenth century. In the 1830s and 1840s, as its object shifted from the state to society and economy, it was more and more limited to social numbers. By the 1850s, however, a few were saying that it was properly defined by its reliance on numerical methods, which happened to apply especially well to the science of society. This emerging sense of statistics as a form of quantitative reasoning points to the continuity between the mathematical field of statistics and its political and administrative forms. In practice, nineteenth-century statisticians were overwhelmingly associated with bureaucratic agencies for recording population, trade, education, crime, poverty, migration, mortality, and madness. To the end of the century, most statisticians (*statists* in English) were still reluctant to let their enterprise be redefined as an auxiliary science—in German *Hilfswissenschaft*—or aid to other sciences. This tension was never resolved. Since about 1900 there has been a scientific field of statistics, defined mainly in mathematical and methodological terms, overlapping with, but mostly distinct from public or official statistics,

[2] R. Steven Turner, 'The Growth of Professional Research in Prussia, 1818–1848: Causes and Context', *Historical Studies in the Physical Sciences*, 3 (1971), 137–82.

which continues to be carried out principally in state agencies or by scholars concerned with state and economy.[3]

Even this wide formulation ignores some of the most interesting aspects of the statistical sciences. Whole industries have grown up around certain statistical tools. These include insurance, various forms of probabilistic modeling, randomized experiments, industrial quality control, regressions (much favored in econometrics), estimations, and the whole world of social surveys from social-science research to political polling and marketing surveys. And still we barely scratch the surface. Currently there is a move afoot to rebrand statistics as data science, which would be more inclusive and less focused on classically scientific endeavors.

Statistics, as an area of mathematics, has been extraordinarily fruitful for the sciences. It was, however, never an unmoved mover. The concepts and techniques of statistics, from error theory and correlation to stratified sampling and analysis of variance, grew up in interaction with natural and social sciences. In practice, the mathematics has never been cleanly separable from its uses. While researchers certainly do sometimes apply statistical methods in ways that are detached from their own disciplinary affiliations, the usual practice is for graduate students to learn their basic statistical methods from courses and textbooks devoted to their own academic field. Perhaps every important statistical tool or problem has a history reaching back to a time before there was any such thing as mathematical statistics. That is, the history of the creation and transmission of research methods in statistics is a highly variegated one. Only in a very loose sense has it been the story of a discipline.

Learning on the Job

Many of the formal tools of statistics can be linked to techniques of aggregating, correlating, handling residuals, and planning experiments. A wide range of problems like these had arisen already in the early modern period, and if they were not marked off as a specific category of problems, the researchers at least were learning from one another's examples. The method of least squares, developed in astronomy and geodesy to calculate a best value from a swarm of measurements and to estimate the bounds of error, was formally articulated in the first decade of the nineteenth century.

[3] Historical studies of these two sorts of statistics were mostly oblivious to one another for a long time, but are brought together in different ways in Theodore M. Porter, *The Rise of Statistical Thinking, 1820–1900* (Princeton, 1986); Ian Hacking, *The Taming of Chance* (Cambridge, 1990), and Alain Desrosières, *The Politics of Large Numbers* (1993), trans. Camille Naish (Cambridge, 1998). There is by now an extensive scholarship on these questions.

Before long there were textbooks, and least squares became as fundamental as telescopes and thermometers to the work of observatories.[4] The practices of social statistics were less precisely articulated and were communicated mostly in a less formal way. Government bureaus concerned with population, trade, health, crime, and the like typically put out numbers without revealing much about their methods. The work, however, was labor-intensive, and special forms of expertise inevitably developed within the offices. For example, medical statistics on the results of smallpox inoculation and vaccination were recorded and shared within networks of doctors, who also discussed and debated their methods. Life insurance actuaries, some of whom made astronomical observations in their spare time, shared data techniques with one another and eventually organized actuarial societies. The Statistical Society of London, founded in 1834, provided a meeting place and a journal for a variety of statistical compilations. It also was a model, perhaps unneeded, for related organizations at home and abroad, some bearing the name *statistics*, others not. An American Statistical Association was organized in Boston in 1839 and, like the English society, has a continuous history up to the present. Until at least 1870, such organizations were much more interested in getting data to guide social reform than they were in working out methodologies of statistical reasoning. But often enough they did not agree, and dissent provided an excellent stimulus to rouse these social quantifiers from their empiricist slumbers.[5]

As early as 1785, M. J. A. N. Condorcet and Pierre-Simon Laplace deployed serious mathematics to calculate probabilities of correct judicial decisions in relation to the size of the jury, on the assumption of a fixed probability that each juror would decide correctly. Siméon Denis Poisson continued the work in the 1830s using official data from French courts. The Belgian observatory director Adolphe Quetelet was almost unique among state statisticians in seeking to understand tabulated numbers of births, crimes, and marriages in relation to mathematical probability. His 1835 book *On Man*, subtitled *Essay on Social Physics*, achieved a considerable reputation, especially for its insistence on natural laws of social behavior and in relation to questions about human free will. He also took a lead role

[4] Stephen M. Stigler, *The Seven Pillars of Statistical Wisdom* (Cambridge, 2016); Stigler, *The History of Statistics: The Measurement of Uncertainty to 1900* (Cambridge, 1986). See also James Franklin, *The Science of Conjecture: Evidence and Probability before Pascal* (Baltimore, 2001).

[5] Andrea Rusnock, *Vital Accounts: Quantifying Health and Population in Eighteenth-Century England and France* (Cambridge, 2002); Michael J. Cullen, *The Statistical Movement in Early Victorian Britain* (Hassocks, 1975).

in the organization of Belgian statistics, both as savant and administrator. He construed the statistical bureau as a social observatory.[6]

Quetelet had in mind the enlightenment of the public as well as effective state management. The statistical office and the observatory were for him, nodes in an apparatus of quantitative research. He wanted to make the academy into a site of collective research, focusing on periodic phenomena. His topics ranged from motions of the planets, seasons, and blooming times of plants to cycles of human activity as revealed by statistics of birth, death, crime, and suicide. It was definitely statistical, at least in the anachronistic sense of being based on abundant data collection. Quetelet's effort to make his academy into an instrument of the research he favored required that it function also as a training ground for quantitative science. This point emerges clearly in his *éloges* for deceased Belgian academicians, summed up in the history he wrote of 'mathematical sciences' at the Belgian Academy.[7]

These efforts, however, were less about statistics as a distinct field of knowledge than about a style of research that extended well beyond it. Medical statistics was more amenable to systematic application as expert knowledge. Jules Gavarret in 1840 had applied Poisson's basic formula to determine whether the difference in outcomes associated with a treatment under investigation could with sufficient assurance be attributed to its genuine efficacy rather than to random fluctuation. A considerable number of German doctors, most of them practicing in insane asylums, subsequently used this formula of Poisson's. While this seems to attest to the openness of these doctors to basic probability theory, I am aware of no evidence that it was ever taught as part of a medical curriculum.[8]

There were professors of *Statistik* in Germany going back to the eighteenth century. While it stretches things somewhat to call this study a discipline, it was taught in universities, sometimes under its own name and sometimes as an aspect of cameralism—the study of economic affairs in relation to the management of state budgets. A related study, political economy, appeared there in the early nineteenth century as a field of study, typically as an alternative and rival to law as the study best fitted to the formation of state officials. Economic study typically included some

[6] Charles Gillispie, 'Probability and Politics: Laplace, Condorcet, and Turgot', *Proceedings of the American Philosophical Society*, 116 (1972), 1–20; Loraine Daston, *Classical Probability in the Enlightenment* (Princeton, 1988); L. A. J. Quetelet, *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale* (Paris, 1935). Joseph Fourier had earlier written in a census volume on the mathematics of population; see Porter, *Rise of Statistical Thinking*, 97–8.

[7] Quetelet, *Histoire des sciences mathématiques et physiques chez les Belges* (Brussels, 1864); Porter, *Rise of Statistical Thinking*, 40–55.

[8] J. Rosser Matthews, *Quantification and the Quest for Medical Certainty* (Princeton, 1995); Theodore M. Porter, *Genetics in the Madhouse: The Unknown History of Human Heredity* (Princeton, 2018), 185, 302.

statistics, and both fields were sometimes included as topics within *Staatswissenschaft*, state science, which appeared on the scene under that name in the 1820s. These complex and rather fussy details are of secondary importance here, apart from the general point that statistics was being taught at universities, mostly for pragmatic reasons and without depending on a clear disciplinary status. The institutional and historical form of political economy that achieved dominance in Germany by 1870, mainly under the name *Nationalökonomie*, was as devoted to statistics as to history, and dismissed the individualism of English political economy as merciless *Manchestertum*. By then, economy had a clear disciplinary status, including many university chairs, yet its orientation in Germany was pragmatic and applied rather than scientific.[9]

The first systematic program of statistical education arose in a similar milieu, yet one still more bound up with official, administrative statistics. Ernst Engel, who came to Berlin from Saxony, had been trained in chemistry and mining. One of his first actions as head of Prussian statistics was to negotiate the creation of a Statistical Seminar. It began in 1862 with just eight students, but by 1872, in the aftermath of German unification, it had grown to 32 students. The course was designed for statisticians within the Prussian state, and subsequently in other German ones, and was part of a strategy to upgrade and harmonize statistical procedures. His own revision of techniques for taking a census and sorting and tallying the results was much admired in Prussia and beyond. His reform effort began with a shift from registering families in books, one line per family, to recording each individual separately on a data card. The new system greatly facilitated the process of sorting individuals and converting the results into diverse forms of tables, sometimes with several variables along the rows and columns of a single table. One very practical aim of his seminar was to bring statistical practices in different ministries into line with the census office. He tolerated, to a degree, visitors from outside of Germany, but the seminar was not mainly for them. While it was connected to university teaching in the state sciences, the seminar was for civil servants, not university students—at least not until they crossed over the line from university to state administration.[10]

[9] Andre Wakefield, *The Disordered Police State: German Cameralism as Science and Practice* (Chicago, 2009) presents a cynical view of the eighteenth-century cameralists; David F. Lindenfeld, *The Practical Imagination: The German Sciences of State in the Nineteenth Century* (Chicago, 1997) give a much more appreciative view of the nineteenth.

[10] Michael C. Schneider, *Wissensproduktion im Staat: Das königlich preußische statistische Bureau, 1860–1914* (Frankfurt, 2013), 131–56; Morgane Labbé, 'Institutionalizing the Statistics of Nationality in Prussia in the 19th Century', *Centaurus*, 49 (2007), 289–306; Christine von Oertzen, 'Machineries of Data Power: Manual versus Mechanical Census Compilation in Nineteenth-Century Europe', in Elena Aronova, Christine von Oertzen, and David Sepkoski (eds.), *Data Histories*, *Osiris*, 32 (Chicago, 2017).

Engel liked to compare his statistical seminar with a chemistry laboratory. It would be rash to dismiss his claim out of hand. The work of the census, especially under Engel, was closely tied to economic and social research on vital topics such as ethnicity, labor, poverty, health, and social insurance. The scholarship on Engel's seminar, unfortunately, includes much more about his negotiations to set up administrative and funding arrangements than on what sort of educational program he developed. Georg Knapp, nephew of the great chemist Justus Liebig, was among the rather few statisticians of his era who had mathematical training. It is possible to trace the evolution of a mathematical form of state statistics in central Europe and Russia, often involving the movement of mathematicians into this social field.[11] Engel's seminar, however, was not a site of mathematics. Reflecting back on his own experience in the seminar in 1865 and 1866, Knapp recalled approvingly the social instruction, and at the same time spoke mockingly of students for whom a simple logarithm was treated as if secured by seven seals. Yet it is evident that Engel's statistical seminar belonged to the culture of the German university, whose commitment to research and to science could assume a wide range of forms.[12]

Statistics Was a British Science? Biometry and Statistical Mathematics

The nineteenth-century predecessors of what in the following century became a mathematical field of statistics appears idiosyncratic to the point that no explanation in terms of broad disciplinary developments seems at all promising. Quetelet wanted to see the astronomical and meteorological work of the observatory integrated with census tallies and with tables of social phenomena such as crimes and suicides. Francis Galton worked on his own for decades on the presumed transmission within families of exceptional talents, and subsequently on statistical patterns of inherited size in peas and then people. Karl Pearson set out to build a discipline in a way that his predecessors had not, but he, too, depended on assembling a statistical edifice out of highly disparate elements. It appears quite different from the systematic programs of disciplinary training that were so successful in German philology, chemistry, and mathematics.

[11] Porter, *Rise of Statistical Thinking*, chap. 8; Martine Mespoulet, *Statistique et revolution en Russie: Un compromise impossible (1880–1930)* (Rennes, 2001).

[12] Schneider, *Wissensproduktion*, 131–2; Georg Friedrich Knapp, *Aus der Jugend eines deutschen Gelehrten* (Stuttgart, 1927), 154.

It could well be that the less consolidated disciplinary structures of Britain brought more advantages than liabilities for building up a new field of statistical calculation, modeling, and inference. As it happens, the most mathematical among German statisticians, such as Wilhelm Lexis and, later, Wilhelm Weinberg, also worked in a relatively independent way. At the same time, none of these statisticians-in-the-making, not even the English ones, were truly independent. Instead, they drew on the materials and formulations of diverse scientific inquirers working in such fields as medicine, agriculture, insurance, and psychology (or psychophysics), each with its own statistical practices. Researches on these topics had considerable value for the emerging biometric school. Galton already was recruiting allies and mobilizing data from experts in these fields in the 1870s. Around 1900, as Pearson rose to prominence, specialists on mental illness, learning disabilities, criminality, and the like quickly recognized the significance of his work for what they were doing. Many took the initiative to contact him, sometimes even before he had learned of their data and expertise.[13]

Still, the English biometricians played a crucial role in shaping statistics as a mathematical field. Even Ronald Aylmer Fisher, who took his undergraduate degree in 1912, faced a world with no established curriculum and no recognized career track for a statistician. It cannot be a coincidence that Galton, Pearson, and Fisher all studied at Cambridge University. All three underwent an intense training oriented around a celebrated mathematical competition, the Tripos. From the standpoint of a mathematician on the European continent, Tripos mathematics seemed more like mathematical physics or applied mathematics. It worked very well, however, as the basis for a career in statistical mathematical sciences. Its focus was not on rigorous proofs, but on solving problems. Anyone who hoped to have a chance of excelling in this competition had to sign up with an experienced 'coach', who drilled the students relentlessly on material relevant to the exam. There was some consistency of style over the seven decades separating Galton's study at Cambridge from Fisher's, even if the specific content was transformed almost completely.[14]

Their mathematical strengths were quite different. Galton suffered a breakdown at Cambridge after driving himself relentlessly in preparing for the 'Little Go' or preliminary exercise. In consequence, he never advanced very far in mathematics, and did not even sit for the Tripos. Yet

[13] See Porter, *Rise of Statistical Thinking* on Galton and Porter, *Unknown History*, chap. 10 on Pearson's biometric allies.

[14] Andrew Warwick, *Masters of Theory: Cambridge and the Rise of Mathematical Physics* (Chicago, 2001).

he was extraordinarily creative and had an excellent ability to discern the mathematical structure of a scientific problem whose solution was beyond his powers. He understood enough, however, to describe the problem so that a trained mathematician could derive a solution. Galton also worked very skillfully with visual representations, most notably a mechanical one, his 'quincunx', a diagonal matrix of pins on a board through which little balls of shot fell and rebounded. At the bottom would appear one or more bell-shaped normal curves of variable width. The quincunx served him as a model of statistical variability, shaped by processes of reproduction and selection.

Pearson, a remarkable social and historical visionary, and at times a bold if unsuccessful physical theorist, achieved an impressive technical competence in mathematics, an ability to set up complex algebraic problems and press forward to a solution. He devoted great effort to fitting curves, and he spared himself no trouble in working out the tangled effects of reciprocal correlations in heredity. He envisioned a world made impersonally efficient by means of scientific method in the form of statistics. Fisher's training in statistics began with error theory and the method of least squares, then extended to evolution and eugenics. He was better able than Pearson to cut through swarms of algebraic symbols to achieve an elegant reframing of a statistical problem.[15]

Galton, who lived from an inherited fortune, marched to his own drummer. By the 1890s, he had come to see statistics as potentially a distinct methodological field devoted to reasoning about empirical numbers and measures. By this time he was especially caught up in the study of evolution and biological inheritance, but he was also coming to realize that some of the relationships he had at first understood as biological principles were more general than that, and could be applied to data from any field. His decisive moment in this regard came at the end of 1888 when he worked out the basic geometry of correlation. About then he began working to encourage young mathematicians to devote their careers to the mathematics of statistics. Although he never really acted as a teacher, he corresponded with younger men and made suggestions. His ideas were picked up in several countries, but especially at home in Britain. His most devoted admirer, and the most important for statistics, was Pearson, who, around 1895, took up statistical mathematics as his great intellectual cause. Like Galton, he was especially impressed by its potential importance for

[15] Stephen M. Stigler, 'Darwin, Galton, and the Statistical Enlightenment', *Journal of the Royal Statistical Society, Series A*, 173 (2010), 469–82; Theodore M. Porter, *Karl Pearson: The Scientific Life in a Statistical Age* (Princeton, 2004).

evolution and eugenics, but his ambitions knew no limits, and he intervened in the work of diverse disciplines, as he later boasted, like a buccaneer.

Pearson's intellectual range and competence were dazzlingly wide. Galton confirmed and to some degree redirected his interests in quantitative reasoning, but certainly was not responsible for Pearson's initial interest in statistics. That came, interestingly enough, from his work as a college teacher of applied mathematics. After several years of uncertainty as to where his best talents lay and where he could make a difference, he accepted a position at University College, London, in applied mathematics. Immediately he set to work to redefine engineering there as a field rooted in mathematics and measurement. In contrast to the mathematical students he knew from Cambridge, the engineering students at UCL put little faith in abstract science. He complained of their preference for working with their hands over devoting themselves to the acquisition of effective intellectual skills. The 'engineering laboratory' that provided the principal focus of his teaching was not designed to turn engineers into mathematicians, but focused on 'graphical statics', for the solution of practical mathematical problems. These techniques, growing out of engineering traditions from the period of the French Revolution, had been developed for the instruction of engineers mainly in Italy and Germany. Pearson was not content to defend these techniques as within the reach of imperfectly-educated engineers, but also, and principally, as a way of making mathematical reasoning visual and intuitive, as it had not been since the triumph of algebra (and analytic geometry) more than two centuries earlier. He set about developing graphical methods of statistics as an offshoot of these engineering initiatives, for the sake of lectures he delivered at just this moment to commercial students at Gresham College in the City of London.[16]

Pearson's philosophical book on science, which acquired a cult status in certain circles during the early decades of the twentieth century, emphasized the moral and political virtues of scientific method. The first edition of this book, which began as another set of Gresham Lectures, was completed just before he turned to statistics as his life mission. Many of his claims there for scientific method seem to resonate with his emerging view of statistics as the all-purpose instrument of scientific reason. From his youth, he had spoken often of an alliance, almost an identity, of science and socialism. Scientific method, which here referred chiefly to something on the order of the scientific spirit, required a person to accept as true only what can be held

[16] Porter, *Karl Pearson*, chap. 8; and Pearson, 'Contributions to the Mathematical Theory of Evolution, II: Skew Variation', *Philosophical Transactions of the Royal Society of London* (A) 186 (1895), 343–414, Fig. 18.

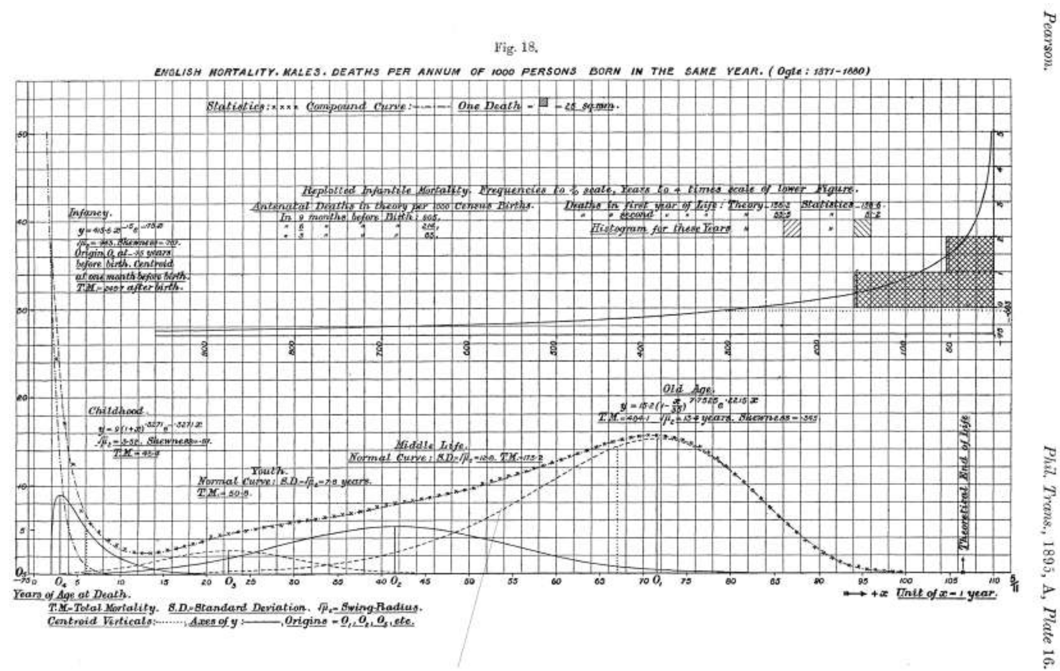

Figure 7.1 An example of Pearson's graphic method. As he turned from graphical statics in engineering to statistics, Pearson focused increasingly on curve-fitting, mainly to approximate the data but also to clarify causes. The most memorable of these efforts was the curve he drew to fit annual death rates of English males, an irregular curve with one peak near the age of birth and another at about age 70. His solution was a sum of five skew curves from a family of curves that he had just developed. In a popular lecture, he explained how the curve of mortality in relation to age was as if produced by five marksmen, shooting with widely varying degrees of accuracy at travelers of a particular age crossing the bridge of life. Pearson, 'Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material', *Philosophical Transactions of the Royal Society of London, A*, 186 (1895), 414.

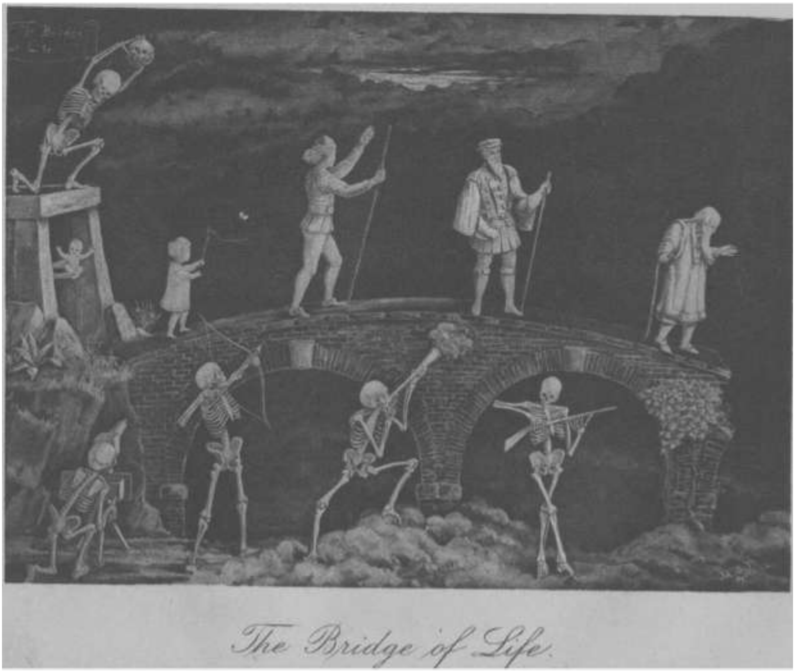

Figure 7.2 Pearson's wife Maria Sharpe created a striking image to illustrate this process. 'The Bridge of Life,' Frontispiece to Pearson, *The Chances of Death and Other Studies in Evolution* (London, 1897).

disinterestedly because it is valid for everyone. Statistics, as a form of reasoning, seems like the fulfillment of a moral commitment to take full account of the evidence, even to reduce reason to calculation. His own theory of knowledge was at stake, in a way, in his efforts to transmit this method of analysis and evaluation to the public, and in particular to his students.

Pearson, however, never supposed that statistics should make intellectual and moral decisions routine or mechanical. He insisted on the contrary view, idealizing the intimate relationship of master and apprentice in the medieval university as a living model for science, still. Although he gave lectures at the most advanced level as well as to undergraduates, he refused to write a textbook or even to teach from one. A student must not be satisfied to learn rules, but must mature into a comprehension of the craft. Students and assistants in his statistical and eugenic laboratories recall 'the professor' making his daily rounds to discuss the research of every student and colleague. They were impressed and inspired by these interactions, and at the same time were oppressed by them. Some of the women and most of the men in his statistics laboratory had a falling out with him at some point. They felt the weight of his brilliance and of his stubbornness. It was difficult to confine these disagreements to the intellectual and technical dimension. He accused them of disloyalty, of abandoning the sacred statistical cause. Some of the most able felt compelled to break with him for a time. Such deep and disturbing disagreements might be dismissed as reflecting Pearson's personal characteristics, which they did. Yet he was not necessarily always in the wrong, and statistics, over the twentieth century, was subject to a series of deep divides, sometimes boiling over into bitter controversies. Some have been impossible to resolve.[17]

Udny Yule and Major Greenwood, though relatively long-term associates of Pearson's laboratory, never held permanent positions there. A succession of young men and a few women came to Pearson's lab as paid researchers, with the expectation of moving on after a few years. They collaborated in his research projects and wrote some papers on their own, and often moved on to jobs that made use of their statistical skills or even to carry out statistical research, despite the failure thus far of statistics to achieve recognition as a job category. Another group of researchers, all women, held positions in Pearson's labs that stretched out for decades. They enjoyed less independence and, in general, received lower salaries, but Pearson encouraged them to coauthor research papers and even to publish independently. His primary eugenic project was more medical than statistical, yet his laboratory associates became expert also in statistics. It is necessary to understand that most of their time, and even of Pearson's, was devoted to the procurement of relevant data and to putting it into an appropriate shape for statistical study. Before the skulls could be analyzed and classified, they had to be measured along many dimensions with calipers, then photographed, perhaps, from multiple angles. School assessments had to be compared with bodily measurements and medical assessments. After that came days and weeks with a Brunsviga calculator to process the data. Pearson, too, kept one always at his side. The need for appropriate formulas to analyze data was unquestioned, yet mathematics here was the tip of the iceberg. Much of the rest was data work.[18]

[17] Porter, *Karl Pearson*, chap. 9; on statistical controversies see Donald MacKenzie, *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge* (Edinburgh, 1981).

[18] Rosaleen Love, "Alice in Eugenics-Land': Feminism and Eugenics in the Scientific Careers of Alice Lee and Ethel Elderton', *Annals of Science*, 36 (1979), 145–58; Eileen Magnello, 'The Non-Correlation of Biometrics and Eugenics: Rival Forms of Laboratory Work in Karl Pearson's Career at University College London', *History of Science*, 37 (1999), 79–106, 123–50.

For many years, Pearson brought in research students, typically for a year, to study in his lab and to collaborate in the research. They came from diverse backgrounds and from several countries, sometimes from other continents. None of them arrived as a statistician in Pearson's sense, since no such field existed. Instead, he taught economists, psychologists, biologists, anthropologists, criminologists, medical statisticians, and, in an especially well-known case, a brewer, W. S. Gossett, who in the course of his visit develop a new statistical test. Pearson took a dim view of government statisticians, the sorts of people who directed census offices, and he refused to have anything to do with their organization, the International Statistical Institute. He allowed them, however, to visit his laboratory and to learn its techniques. One of the most successful of these visitors was Prasanta Mahalanobis, who had a lead role in establishing the Indian Statistical Institute in Calcutta, and eventually became a leader of economic planning in independent India.

Experiment and Inference

Like Pearson, Fisher was deeply committed to the quantitative study of heredity. Both had important roles in the articulation and defense of Darwinian evolution, and each was outspoken on the urgency of eugenic research. While Pearson's hereditary studies dealt mainly with quantifiable traits, Fisher wanted to get beneath the traits and apply his statistics to the presumed Mendelian factors. Genetics, in alliance with eugenics, was one of the most important fields of application for statistical methods. It was, more than that, of crucial importance for the articulation of statistical mathematics, just as biological, medical, and anthropological studies provided the most important topics for much of Pearson's work. Fisher, like Pearson, took an active interest in several distinct scientific fields. His statistics, though intimately bound up with genetics, owed no less to the stimulus of agricultural research. It was above all his work that reshaped the role and identity of the statistician.

Pearson and Fisher, after some tense but respectful early interactions, became bitter antagonists. When, in 1919, Fisher had the opportunity to take up a research position in Pearson's lab, he chose to keep clear of the constraints that Pearson would impose, and accepted instead a post with the agricultural station of Rothamsted, just north of London. Agriculture was a familiar topic of statistics, most obviously in the form of crop summaries, but more profoundly as a focus of controlled experimentation in order to increase of crop yields. Fisher conceptualized the problem in a new way, based on individual plots as the unit of analysis. These plots were to be compared with control plots, always on the basis of adequate replication, and the choice of experimental and control plot would be based on a procedure of randomization, in effect, by tossing a coin. The specific purpose of the exercise was to achieve sufficient confidence that the measured difference of a new fertilizer, for example, was genuine, that is, not the result of meaningless fluctuation. Randomizing brought the experiment into comformity with basic assumptions of probability theory, permitting the calculation of what he called the likelihood that such a difference might have arisen merely by chance. If this likelihood was sufficiently low, for example, below 5%, the difference between treatment and control would, provisionally, be taken to be real.

This procedure tended to shift the focus of a statistical experiment away from a measure of the strength of a relationship between variables to a probability measure. Significance in this context is statistical, not substantive, referring not to the importance of the effect, but to the confidence that the effect is nonzero. Although Fisher allowed that different significance levels might be appropriate in different circumstances, most disciplines fixed on a particular one, 0.05, with the more strenuous 0.01 as runner up. On occasion he sharply criticized the fixation of researchers on particular significance levels, yet he declared in 1935 in his authoritative book on *The Design of Experiments* that "every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis.' This kind of testing became the heart of inferential statistics, what many social and natural scientists construed as the essence of scientific methodology. Fisher here offered a vision of scientific inference as technically demanding, perhaps, but highly routinized, preferring a clear standard to a result that matters.[19]

Agricultural researchers, already familiar with statistics in several forms, were not slow to recognize the promise and coherence of Fisher's methods. Before long, students began coming to Rothamsted. His agricultural experiments became known internationally, leading eventually to invitations to visit US agricultural schools at Iowa State College and North Carolina State, both of which became, in turn, important centers of statistical research and the diffusion of statistical methods. *The Design of Experiments* showed that his alliance of experimentation and experiment extended far beyond agriculture. It begins with the homely example of a lady who says she can tell whether her milk has been put in before or after the tea. Ian Hacking pointed to a resemblance between the lady's claim and those of psychical research, arguing that the technique of randomization

[19] Gerd Gigerenzer, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, Lorenz Krüger, *The Empire of Chance: How Probability Changed Science and Everyday Life* (Cambridge, 1989), Fisher's quote from 211.

may have originated in psychical research. Trudy Dehue then explained how randomized controls had entered and become routine in the psychology of learning still earlier. The absolute priority is perhaps not important, especially since the forms and purposes of randomization were far from uniform. If psychologists did not articulate a developed strategy of randomization in advance of Fisher's, they were in an excellent position to notice the statistical mode of experimentation and to systematize it for their own discipline. Farm plots were replaced by schoolrooms and other laboratory spaces, and fertilizer by a curricular modification or the introduction of a preliminary exercise. A rigorous and impersonal methodology, which might be important to persuade schools or militaries to take seriously the claims of school reformers, should at the same time enhance the scientific reputation of the discipline.[20]

Applied researchers and social scientists began looking to statistics for a purely objective scientific standard. This depended on closing their eyes to the bitter controversies by which the field was riven. Karl Pearson's rejection of Fisher's statistical program was perhaps losing its credibility by 1920, but the Polish immigrant Jerzy Neyman, who received some of his training in Pearson's lab, teamed up with Pearson's son Egon to frame a different program for statistics that was at odds with Fisher's significance testing. Gerd Gigerenzer showed how writers in psychology defined a new statistics, bringing together as needed pieces that both Fisher and Neyman regarded as incompatible. We see here how the usual assumption about hierarchies within science breaks down. Training in statistics altered the basic character of the psychological experiment, just as it had reshaped the agricultural one. At the same time, fields of application like agriculture and psychology reinterpreted and reshaped what was being worked out as a new mathematical field.[21]

Disciplines and Professions

To some degree in the 1930s, and then with a vengeance after the Second World War, the new tools of error management and of statistical inference provided a revised basis for the human sciences, therapeutic medicine as

[20] Gigerenzer et al., *Empire of Chance*, chap. 3; Ian Hacking, 'Telepathy: Origins of Randomization in Experimental Design', *Isis*, 79 (1988), 427–51; Trudy Dehue, 'Deception, Efficiency, and Random Groups: Psychology and the Gradual Origination of the Random Group Design', *Isis*, 88 (1997), 653–73; John Carson, *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750–1940* (Princeton, 2007).

[21] Gerd Gigerenzer, 'Probabilistic Thinking and the Fight against Subjectivity', in Lorenz Krüger, Gerd Gigerenzer, and Mary S. Morgan, (eds.), *The Probabilistic Revolution, ii: Ideas in the Sciences* (Cambridge, 1987), 11–33.

well as social science, and for other fields as well. Statistics was identified with objectivity and with the rigorous neutrality that science was taken to demand and now seemed to be able to supply. The social disciplines were particularly insistent on drawing a sharp line between real social science and the well-meaning efforts of soft-hearted social reformers. This rejection of moralizing language, however, did not prevent social science from offering guidance to policy initiatives, which they preferred to express using terms like adjustment or efficiency, helping the social system to function smoothly, and getting more bang for the buck. Statistics, as we have seen, had long been associated with broadly professional and administrative activities pertaining to industry, agriculture, schooling, health, housing, and poor relief. Problems of classification associated with mandatory schooling stimulated the introduction of new statistical techniques into psychology. New statistical tools, including a new economic field, econometrics, developed in response to interwar economic instability. Statistics still did not mean just one thing; the more engaged forms of social science included different sorts of statistical tools from those demanding academic rigor above all else.[22]

In psychology, too, the most basic statistical tools had been framed in contexts of application even before these fields became university disciplines. Although they were responsive to developments in the new mathematical field of statistics, the methods they taught remained in an important sense their own. Public health was like this too, the product of a long tradition that emphasized environmental causes of sickness and mortality and that carefully tracked the progress of epidemics. The economy of clinical medicine, by contrast, was anchored in individualized relationships between physicians and their paying patients, and the training of clinicians had usually emphasized this dimension. Although the therapeutic trial did not arrive out of the blue, it depended a good deal on the initiative of regulators and statisticians. Statistical medicine was as important for Karl Pearson and his students as were evolution and eugenics. The randomized trial, which had multiple sources, was eagerly taken up psychologists and more hesitantly by clinicians. Even for them, the introduction of the RCT does not come down to the passing of a baton from one great statistician to another. Instead, a range of pressures and incentives involving medical researchers, regulatory authorities, and

[22] Kurt Danziger, *Constructing the Subject: Historical Origins of Psychological Research* (Cambridge, 1990). On the problem of economic cycles and origins of econometrics, Mary S. Morgan, *The History of Econometric Ideas* (Cambridge, 1990).

pharmaceutical companies stimulated a move toward systematic evaluation of drugs.[23]

Although statisticians did not have to break down the doors of clinical medicine, neither did they draw mainly from their own traditions. This is in contrast to econometrics, which borrowed more heavily from economic traditions than from mathematical statistics, and even from experimental psychology, with its crucial sources in educational studies. Although the social sciences drew heavily from the new mathematical field of statistics, they did so by assimilating it to their own traditions. Medical statisticians, by contrast, typically recruited from outside, sometimes in defiance of the customary individualism of clinical practice. It matters, too, that modern medicine has long been rich enough to import statistical experts to preside over the design of therapeutic experiments. Some became medical specialists in their own right. Doctors participating in large-scale research had to sacrifice the expert discretion that came with medical individualism in order to participate in large-scale clinical trials. Clinical medicine had no clear precedents for this. The logic of the randomized trial came from statistics.[24]

Who is a Statistician?

Medical statistics, too, developed almost immediately into a distinctive form of statistical practice, and soon, this special form of statistics began to be taught in medical schools. In fields like psychology, economics, and ecology, some scientists achieved a level of statistical expertise approaching that of professional statisticians. The founding of the Institute of Mathematical Statistics in 1935 may be taken as a convenient marker for the emergence of mathematical statistics as a partly autonomous field, one that was beginning to train its own students. However, the perpetuation of statistical expertise is much more interesting and complicated than the model of autonomous disciplines would suggest. As statistics became a routine and necessary tool in a wide variety of disciplines and practices, their faculties learned to teach using methods and examples that were often specific to the subject disciplines. Many or most of their students lacked the preparation to take a graduate course taught by a mathematician.

[23] Martin Edwards, *Control and the Therapeutic Trial: Rhetoric and Experimentation in Britain, 1918–48* (New York, 2007), esp. 14.

[24] There is now a considerable literature on the origins of the clinical trial and its politics. The classic work is Harry Marks, *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900–1990* (Cambridge, 1997); see also Gérard Jorland, Annick Opinel, and George Weisz, (eds.), *Body Counts: Medical Quantification in Historical and Sociological Perspective* (Montreal and Kingston, 2005).

Statistics, in short, has never been easy to isolate from other forms of knowledge. Beyond that, it is increasingly regarded as an indispensable component of common knowledge. The work of census offices and other sites of official statistics remains central to this aspect of the field. Right from the start, in the early nineteenth century, reformers and politicians liked to claim that the numbers would speak for themselves. Throughout the history of public statistics, it has been common to argue that statistics can support the interests of the citizenry by revealing in simple, numerical terms if a political program was sound or misguided. By the twentieth century, it was no longer only the critics who argued that valid and informative numbers might require the involvement of experts. Statistics emerged in the 1930s and 1940s as a highly dispersed field, taught sometimes as a branch of mathematics, sometimes as a distinct discipline specializing in the management of chance and variability, and sometimes as a workbook of practical techniques for turning the data of a discipline into acceptable research papers.

The tools of probability that Pearson, Fisher, and others began using for biological, agricultural, medical, and social data eventually came back to government statistics. Census offices, commerce ministries, departments of agriculture, and bureaus of labor statistics had not been inert during the decades when it took shape as a mathematical field. Outside the Anglophone world, state statisticians remained the most visible experts in statistical work.[25] They continued to dominate the International Statistical Institute, and they have had a key role in negotiating such unity as the European Union has been able to establish. They already were moving toward a more mathematical and method-conscious form of statistics by the late nineteenth century. They were not slow to take up mathematical tools of analysis, for example, to estimate the bounds of error and to show that an apparent effect could not reasonable be attributed to chance. Official statistics had always been engaged in systematic social observation. By 1900 it was being called on for data on specific matters that seemed relevant to pressing policy issues. Statisticians already recognized that samples should somehow represent the population, and before long began to emphasis the need to represent its heterogeneity. The distinctive value of random sampling was more readily visible to mathematicians, who perhaps also underestimated the difficulty of lining up a fair simulacrum of a large, dispersed population.[26]

[25] See for example Jean-Guy Prévost, *A Total Science: Statistics in Liberal and Fascist Italy* (Montreal and Kingston, 2009).

[26] On sampling in the history of statistics see Desrosières, *Politics of Large Numbers*. Especially good on the social survey is Sarah E. Igo, *The Averaged American: Surveys, Citizens, and the Making of a Mass Public* (Cambridge, 2007).

The tools of statistics arose partly in situations of practice, intersecting with yet never dominated by more academic forms. The latter, of course, had the advantage of their teaching roles--their close contact with students at a moment of relative openness or susceptibility. Later, however, they must face the challenge of implementing an academic program within a community of practitioners who have worked out tricks to elude somehow the rigidity of mathematical theory. Statistics, in short, presents a complex model of academic training for a multifarious occupation, and nothing so simple as the creation of a discipline through systematic university-based research training.

University of California, Los Angeles