In 1990 I had the occasion to review for ISI's *Short Book Reviews* the volume *"Student": A Statistical Biography of William Sealy Gosset*. The book (Pearson 1990) had been begun by Egon Pearson and was completed after Pearson's death by Robin L. Plackett and George A. Barnard. In the course of the review, I commented on Gosset's admirable character and attractive writing style, but ended on a provocative note:

> Gosset possessed excellent statistical insight, and he was surely a catalyst to some important developments in statistics. But there has long been a tendency to exaggerate his achievements, I suspect in recognition of his admirable character, and without a more extensive study it is difficult to judge whether he was an essential catalyst.

As it happened, Frank Yates had been asked to review the book before it came to me but had declined (after a reading that left marginal notes on the copy that came to me). Although Yates, who had known Gosset, may not have found the book sufficiently to his taste to review, he nonetheless rose to my provocation to write a spirited rejoinder, which the editor of *Short Book Reviews* published in a later issue, the only time in 25 years that such a reply appeared. Yates, not surprisingly, thought Gosset was much more than a catalyst, but I still think my observation was a fair one, and I do not pretend to know even now the answer to the question I asked, of whether he was an *essential* catalyst; that is, would the development of statistics have been much different had Gosset never traveled to study in Pearson's laboratory in London?

Of course, I do not mean to raise any doubt about the excellence of Gosset's work or the quality of his mind; Sandy seems to me to capture those quite accurately. But there are examples of people who have written excellent works and who have later attracted renown, but who arguably had little or no impact on the development of our subject. Thomas Bayes is one such example. His article received no serious notice before the twentieth century, and I think a case might be made that had his article never been written, only our modern terminology would be different.

Gosset's case is different. As Sandy explains, whereas the 1908 article was essentially ignored by most statisticians for more than two decades, there was a key exception: Fisher. The article's effect on Fisher was clearly important; it surely played a role in exciting his interest in problems of distribution. Gosset's technique (essentially Pearsonian moment calculation and curve fitting) had no visible effect on Fisher, who immediately adopted a totally different approach; what influence the article had was due to the problem it addressed, not to Gosset's attempted solution, I would suggest. Even if Gosset's guess of the distribution of s^2 had been wrong, I think the effect would have been the same, and the article gave no hint of any of the magic that Fisher produced in the beautiful interrelations among the distributions of the t-statistic, the two-sample t-statistic, the correlation coefficient, regression coefficients, and the sums of squares for analysis of variance.

Fisher's laudatory obituary of Gosset (Fisher 1939) may be read as supporting this view. After four and a half pages of the highest praise for Gosset (and a few digs at Pearson), Fisher took much of it back by stating that he doubted Gosset had understood the full measure what he had done. Indeed, I think Fisher was accurate in that assessment, if we take his description of Gosset's accomplishment at face value. As Alfred North Whitehead wrote in 1917, "Everything of importance has been said before by someone who did not discover it." From this point of view, the importance of the 1908 article is due to what Fisher found there, not to what Gosset placed there. That may have been self-serving on Fisher's part, but I think there is merit in the view, and the fact that no one else found gold in that vein between 1908 and 1922 argues in its favor.

I still view Gosset's primary role as that of an important catalyst. The question of essentiality may be unanswerable, and certainly has no bearing on the decision to celebrate Gosset's achievement of 1908. Gosset was a wise and creative worker who, although thoroughly in the sway of nineteenth century and Pearsonian ideas, wrote one article that caught the eye of the one person who could break free of those constraints. Sandy Zabell's lucid and scholarly analysis of his life and work gives a perfect accent to the occasion of celebrating that article and the man who will always be best known as "Student."

Comment

John ALDRICH

1. INTRODUCTION

When we think of "Student's t," we are at least as likely to be thinking Ronald Fisher's thoughts as Student's. The designation, "t-distribution with $n-1$ degrees of freedom," like the idea of t as one of a family of distributions based on the normal distribution or the application of the t-distribution to regression, were products of Fisher's imagination. That we think of Student's article at all is due largely to Fisher. Professor Zabell quotes Student, writing in 1934, "in the pre Fisher days no one paid the slightest attention to the paper." I would like to develop the theme of Student and Fisher and the "inextricable link" between their histories.

After 1908, Student wrote only three articles on his z distribution; two extended the original tables, and one replied to a criticism that Karl Pearson made in 1931(!). The big advances were made by Fisher, whose studentry can be divided into two phases; in the first phase; he did or redid Student's mathematics, and in the second phase, he took over the z distribution and reconfigured it. The second phase, one of the great passages

John Aldrich is Reader, Division of Economics, University of Southampton, Southampton, SO17 1BJ, U.K. (E-mail: *jca1@soton.ac.uk*).

Journal of the American Statistical Association
March 2008, Vol. 103, No. 481
DOI 10.1198/016214508000000111

in the history of twentieth century statistics, was well treated in the biographies by Box (1978, chap. 5) and E. S. Pearson (1990, chap. 5), whereas Eisenhart (1979) followed the transition from the z of Student (1908a) to the t of Fisher (1925a). Of course, the transition went deeper than just writing $t = z\sqrt{n-1}$. The not so familiar first phase was not so spectacular, but is quite intriguing nonetheless. Fisher (1939) wrote his own record of Student's scientific contribution, but his aim was to instruct, and anything that might distract from the lesson, such as Student's Bayesianism or his own reasons for working on Student's problems, was omitted.

The story will also tell us something about Student the person. The Student–Fisher connection is that rare thing, a sunny story from the history of British statistics in the early twentieth century. Much of the sunshine came from Student, and the story clearly brings out the kind of person he was.

2. THE VERY BEGINNING

Fisher and Gosset first met in September 1922, but they had been corresponding intermittently for 10 years. In 1912, when the first phase of their "collaboration" began, Fisher was 22 years old and Gosset was 36; Karl Pearson was 55. This first phase ended in 1915 with the publication of Fisher's article on the exact distribution of the correlation coefficient; this is the only publication from that time to reveal any connection between the two men's work. They already recognized each other's qualities: "It has been the greatest pleasure and interest to myself to observe with what accuracy 'Student's' insight has led him to the right conclusions" Fisher (1915, p. 508) reports. Student thanked Fisher for "the kind way in which you referred to my unscientific efforts" (Pearson 1968, p. 447). Fisher added mathematical precision to Student's insight; Student set up problems, and Fisher knocked them over. Doing the mathematics for Student was not a small thing, for nobody else could. Zabell identifies one factor behind Fisher's taking on this role—the desire of a young and ambitious mathematician to show what he could do; but Fisher already had his own scientific agenda, and Student's problems happened to be on it.

Fisher's 1915 article grew out of Student's second 1908 article on "the probable error of the correlation coefficient." But, as Zabell relates, there was an earlier nonpublished article based on the first in which Fisher derived the distribution for z. In September 1912, Student forwarded Fisher's derivation to Karl Pearson, suggesting that he publish it. In the covering letter (reproduced in Pearson 1968, p. 446), Student summarized his transactions with Fisher. These began a few months earlier when Fisher sent Student an article that he had written (Fisher 1912) that proposed a new estimation method, the "absolute criterion." Fisher later called this method "maximum likelihood." He applied the method to the problem of estimating the mean and the precision of the normal distribution. Fisher's estimate for the precision (parameterized as $h = 1/\sigma\sqrt{2}$) involved a factor n instead of the $(n-1)$ that was customary in the theory of errors. Fisher criticized some arguments leading to the $(n-1)$ value, but then the argument took an unexpected turn. Fisher mentioned that h could be estimated by choosing the value that maximizes the frequency distribution of the statistic that Student (unknown to Fisher) denoted by s^2. This second procedure actually supplanted the first in Fisher's thinking, for when Fisher (1915) used an analogous procedure to estimate the correlation coefficient by maximizing the frequency distribution of r with respect to ρ, he referred to it as the "absolute criterion." Fisher only gave up this second form of the absolute criterion (for the first) in 1921 (for a more complete account, see Aldrich 1997). Fisher had an interest in obtaining the distribution of s^2, but whether or not he derived it independently of Student (1908a) is unclear; that article must have come up in the course of their correspondence. At the time, Fisher had no real business with the z distribution, and Zabell is probably right that Fisher derived it because he could!

Together, Student's articles of 1908 and Fisher's work of 1912–1915 produced a collection of results in distribution theory, and yet, if I am right, their collaboration was based on different priorities and conflicting approaches to inference. Student was interested in producing a test based on z, and for this the distribution of s was just an input; for Fisher, the distribution of s was wanted for estimation and the distribution of z came as an easy extension. Regarding principles of inference, Student was a Bayesian—of a kind. Zabell (Sec. 3.12) describes the curious Bayesian structure of the correlation article in which a frequency distribution for the correlation coefficient was sought so it could be multiplied by a prior, and also how the z frequency distribution was treated as a posterior without any explicit Bayesian structure to support it. Like Student's thinking about z, Fisher's thinking about the absolute criterion was half-baked, yet there is a clear anti-Bayesian streak in his 1912 article. Student and Fisher seem to have converged on the same problems for unrelated, if not opposed, reasons. To what extent they *exchanged* views on inference is unknown, for only one letter survives—Student's thank you letter for the correlation article. This is the letter to which Zabell refers for Student's pondering the effect of adopting different priors.

In 1908, Student sought exact distributions for three quantities, s, z, and r. The publication of "The Probable Error of a Mean" was *not* a great event. Student had taken the problem to Pearson, who had helped him solve it. Student used Pearson's tools and wrote in his language (see Aldrich 2003), but the problem was not Pearson's, and its solution gave him no cause for celebration. The problem belonged to the Gaussian theory of errors, which Pearson considered defunct. Student's tables went into Pearson's *Tables for Statisticians* (1914); good tables were not to be wasted, even if they were good for very little. Otherwise, Pearson ignored Student's z until 1931. For the biometrics of the time, only the distribution of r really mattered, although Fisher, quixotically, became interested in s as well. The distribution of s was brought up by Fisher in his correlation article (1915, p. 509) and in a postscript to that article, Pearson (1915) added his thoughts on s. Buried in Fisher's work (1915, p. 518) was one new use for Student's z, but Fisher's interest in z caught fire only when he found a use for it in regression in 1922. It was then that the t story took off.

3. THE *t* DISTRIBUTION AND STUDENT

In December 1918 (contact had been reestablished in 1917), Student told Fisher that there might be a job going at Rothamsted Experimental Station: "I don't know whether you are looking for a job in that line, but I hear that Russell intends to get a statistician sometime soon." Guinness had an interest in the cultivation of barley, one of its main inputs, and Student was a figure in the world of agricultural experiments. Fisher was offered the Rothamsted job, and agricultural experiments became his line. E. S. Pearson (1968, p. 448) thought it "very likely" that Fisher's appointment owed something to Gosset's links with the agriculturalists.

In the early years at Rothamsted, Student was Fisher's lifeline to the community of statisticians. He was the first to be told of new results, and Fisher considered him the only person who understood his work. More than 60 letters survive from the period 1922–1925, nearly all from Student to Fisher. A request in the letter of April 3, 1922 (letter 5 in McMullen 1970) precipitated the second phase of Fisher's studentry. Professor Zabell quoted from that letter when describing Student's views on Bayes. Student had seen the first of Fisher's articles on "likelihood" and knew that Fisher rejected the Bayesian argument. He also had seen the first of the articles reconstructing Pearson's chi-squared theory; not only was Fisher discussing significance tests in earnest, but he also had introduced the notion of "degrees of freedom." Gosset wrote: "I want to know what is the frequency distribution of $r\sigma_x/\sigma_y$ for small samples, in my work I want that more than the r distribution now happily solved." From the notation and reference to r, Student was clearly after the solution to another problem associated with the bivariate normal, the distribution of the regression coefficient. Fisher's response surprised him, because it involved relocating the problem from the theory of correlation to the theory of errors; the change was discussed by Aldrich (2005). The solution involved a suitably constructed z statistic, which pleased Student, although he was not easily persuaded that the solution was correct.

A year later, in May 1923, Fisher was reporting an advance of a different kind, an account of the interrelationship between the various distributions associated with the normal distribution. The letter (which appears in Box 1978, p. 118) formed the basis of the synthesis, "On a Distribution Yielding the Error Functions of Several Well-Known Statistics" Fisher (1924), in which Pearson's chi-squared and Student's t distributions (for the first time) appear as special cases of a general distribution that Fisher called z; transformed, this became the modern F $(= e^{2z})$. Amid the general advance, one backward look should be mentioned. The group theorist William Burnside (1923) published a treatment of the Bayesian version of the problem of the probable error of the mean; Pfanzagl and Sheynin (1996) described this work. Fisher wrote a note (1923) registering Student's priority and giving a derivation of z on the lines, presumably, of the rejected piece from 1912. Fisher also provided a clear statement of the difference between the Bayesian and sampling theory projects. When Fisher sent Student Burnside's paper and his own note, Student simply commented, "It is interesting to see how à priori probability has got him just off the line." (letter 25 in McMullen 1970). There was no degrees of freedom adjustment, as it would be called later. In a later letter (letter 39), Student referred to the "futility of à priori assumptions." It is interesting to speculate on how he would have reacted to the argument of Jeffreys (1931), which gave a distribution exactly on the line (i.e., a t with the right number of degrees of freedom); whether he ever saw it is not known.

The new t was proclaimed in two works in 1925. "Applications of 'Student's' Distribution" provides the theory of the applications, and *Statistical Methods for Research Workers* demonstrates the applications. The book would make both Fisher's and Student's names. It is largely a book of three distributions, Fisher's z for the analysis of variance and two from others, Pearson's chi-squared and Student's t. Suddenly the occasional contributor of minor pieces to *Biometrika* was on the pedestal with the master, and, more than that, *his* contribution contained no "serious error." In the fourth edition Fisher (1932, p. 24) wrote that "from the first edition it has been one of the chief purposes of this book to make better known the effect of [Student's] researches, and of the mathematical work consequent upon them."

Fisher also reworked Student's old examples. Zabell (Secs. 3.2 and 3.5) notes how Fisher used one of Student's data sets, the Cushny–Peebles data, to illustrate the t-test. Naturally, Fisher (1925a, p. 108) stripped away the Bayesian language; instead of saying that "the odds are about 666 to 1 that 2 is the better soporific," Fisher concluded from the t value of 4.06 that "only one value in a hundred will exceed 3.250 by chance so the difference between the results is clearly significant." If Student did not like this reformulation, he had at least two opportunities to say so. He read the proofs of the book as a favor to Fisher and reviewed the published work (Student 1926); on neither occasion did he comment on Fisher's handling of Student's distribution. He was not afraid of registering disagreement; he was always skeptical about the use of controlled randomization in experimental design. At the proof stage (letter 50), he commented, "you would want a large lunatic asylum for the operators who are apt to make mistakes enough even at present." He made this point more decorously in the review.

Student's "new tables" of 1925 were for t. Student (1925, p. 105) saw two defects in the existing tables: "as n increases, the z scale becomes very coarse" and "except in the case for which it was designed, n, the number in the sample, is not the best number under which to enter the table, but $n - 1$ the number of degrees of freedom." Student deferred to Fisher in mathematics—in letter 36 he refers to his "Watsoning" to Fisher's Holmes—and he saw Fisher's replacement of z by t as a mathematical advance. But beyond the mathematics, Student's final statement on z is fully Fisherized and entirely de-Bayesed. To correct Karl Pearson's misunderstanding of the z-test (Pearson never acknowledged the existence of t), Student (1931, p. 408) spelled out "what we actually ask ourselves":

> If the average difference between A and B in the population were zero, what would be the probability of obtaining a sample of differences giving a value of z as high as that observed? and if this probability is sufficiently small we say that the difference is significant.

4. FISHER AND STUDENT

In the early years, Fisher received valuable support from Student, the one established statistician who believed in him, and Fisher (1939, p. 8) acknowledged a "loyal and generous friend." Student the man did not need Fisher's help, but to Student the scientist ("one of the most original minds in contemporary science"), Fisher was very generous. Fisher (1939, pp. 5–6) described how he had solved a problem that "the very brilliant mathematicians who have studied the Theory of Errors" had overlooked and worked against the indifference of "the leading authorities in English statistics." Fisher (1939, p. 5) did acknowledge, however, that

It is doubtful if "Student" ever realized the full extent of his contribution to the Theory of Errors. From correspondence with him before the War... I should form a confident judgement that at that time he certainly did not see how big a thing he had done.

The same could be said of Fisher at the same time, but even when *Statistical Methods for Research Workers* appeared, Student did not realize how big a thing he was part of. He (1926, p. 148) welcomed the book that would change statistics as the first book to present the "special technique" required for dealing with small samples.

Had Student read his scientific obituary, he may well have said "oh, that's nothing—Fisher would have discovered it all anyway." That, according to Cunliffe (1976, p. 4), was his response on being thanked rather grandly for all he had done for "the advancement of statistics." Whether there was such an encounter, the tale is true to Student's dislike of pomposity, his modesty, his recognition of Fisher's genius, and his ease with it. The tale also brings out his realism, for Fisher would probably have discovered it all! From the start, Fisher was extraordinarily self-propelled and it is easy to argue that Student's intellectual impact was not on Fisher, but rather on Egon Pearson, who had doubts and discussed them with Student (see Pearson 1990, chap. 6).

The 1912–1925 interactions appear so sunny because there was a dark cloud in the form of Karl Pearson. Fisher saw his own situation as "publish or perish," and every Pearson rejection threatened his existence. He saw Student as a fellow victim, although Student was philosophical about Pearson, and in truth, Fisher could find no greater offense against him than "weighty apathy" toward his writing. Alas, new clouds were visible even in Fisher's memorial to his friend. When Fisher (1939, p. 6) suggested that concern with the "practical interpretation of experimental results" was vital for Student's success, he was taking a swipe not at Laplace or Gauss, but rather at Neyman and the younger Pearson. It became a theme with him that the latecomers had misunderstood his and Student's work—but that's another, and less sunny, story.

ADDITIONAL REFERENCES

Aldrich, J. (1997), "R. A. Fisher and the Making of Maximum Likelihood 1912–22," *Statistical Science*, 12, 162–176.

——— (2003), "The Language of the English Biometric School," *International Statistical Review*, 70, 109–131.

——— (2005), "Fisher and Regression," *Statistical Science*, 20, 401–417.

Cunliffe, S. V. (1976), "Interaction," *Journal of the Royal Statistical Society*, Ser. A, 139, 1–19.

Fisher, R. A. (1925b), "Applications of 'Student's' Distribution," *Metron*, 5, 90–104.

——— (1932), *Statistical Methods for Research Workers* (4th ed.), Edinburgh: Oliver & Boyd.

Pearson, K. (1914), *Tables for Statisticians and Biometricians*, Cambridge, U.K.: Cambridge University Press.

——— (1915), "On the Distribution of the Standard Deviations of Small Samples: Appendix I to Papers by 'Student' and R. A. Fisher," *Biometrika*, 10, 522–529.

Student (1925), "New Tables for Testing the Significance of Observations," *Metron*, 5, 105–108.

——— (1926), Review of *Statistical Methods for Research Workers*, by R. A. Fisher, *Eugenics Review*, 18, 148–150. Available at *http://www.economics.soton.ac.uk/staff/aldrich/fisherguide/student.htm*.

——— (1931), "On the 'z' Test," *Biometrika*, 23, 407–408.

Comment

A. W. F. Edwards

Zabell rightly stresses Student's pathbreaking contribution to statistical thought and practice with his 1908 paper, but let us not forget its literary quality too. The Introduction is a wonderfully clear description of the problem to be solved, the reason why it is important to solve it, and the means by which the author proposes to do so. It is a model of how to begin a scientific paper. The Conclusions at the end are equally clearly stated, and we may note particularly 'Finally I should like to express my thanks to Prof. Karl Pearson, without whose constant advice and criticism this paper could not have been written'.

For the fourth edition of *Statistical Methods for Research Workers* (1932) Fisher added a 'Historical Note' to Chapter I in which he said "'Student's' work was not quickly appreciated, and from the first edition it has been one of the chief purposes of this book to make better known the effect of his researches, and of mathematical work consequent upon them". Incidentally, in 1924 Fisher asked Student to read the proofs of the first edition, and one consequence of this was the incorporation of Student's suggestion that fold-out duplicates of the statistical tables in the book should be added at the end (Edwards, 2005).

It is a mark of the completeness of the revolution in statistical thinking which Student brought about that so little more needs to be said, but Zabell's account of how the mathematical gaps in his argument were later filled, the proofs improved, and the antecedents unearthed, is most welcome. Just one nagging problem remains – fiducial inference, to which Zabell turns in Section 4.4, having already mentioned a 'particularly interesting remark' of Student's in Section 3.2.

Both Zabell and Fisher (1939) have noticed that Student wrote 'if two observations have been made and we have no other information, it is an even chance that the mean of the (normal) population will lie between them', and Fisher remarked that this was an example of a statement of fiducial probability. He went on to note that it could be applied to the median of any distribution, and he generalised the method to samples of any

A. W. F. Edwards is Emeritus Professor of Biometry, University of Cambridge, and a Fellow of Gonville and Caius College, Cambridge, CB2 1TA, U.K. (E-mail: *awfe@cam.ac.uk*).

Journal of the American Statistical Association
March 2008, Vol. 103, No. 481
DOI 10.1198/016214508000000067