

Beyond the Two Cultures: William H. Kruskal on the Importance of Cross-Disciplinary Research

Number 50

December 11, 1989

Readers of *Current Contents*® have frequently been referred to the idea of the "two cultures" made famous by C.P. Snow.¹ Recently I learned of a commentary by my longtime acquaintance William H. Kruskal, Department of Statistics, University of Chicago, Illinois, which uses Snow's paradigm as a starting point. Not content to settle for the simplicity of two cultures, Kruskal borrows a variable from his field of statistics and speaks instead of "*n* cultures."² It is not surprising that, as the purveyor of the *Science Citation Index*® (*SCI*®) and its companions, the *Social Sciences Citation Index*® and the *Arts & Humanities Citation Index*™, I share a predilection for this theme. The *SCI* has from the outset been intended as the ultimate cross-disciplinary index, and I continue to derive endless pleasure from the connections it enables scholars to make. Allowing myself a little self-promotion, I would say that the *SCI* has probably stimulated thousands of intellectual "marriages," and in that sense I have played the broker.

As a discipline, statistics has played an important role in the history of citation indexing. One of the earliest citation indexes, in fact, was developed for statistics and probability.³ It is not coincidental, therefore, that I should have often come in contact not only with Kruskal, but also with the likes of John W. Tukey, Frederick Mosteller, the late James Dolby, and others mentioned in Bill's talk, which was originally presented at the Fourth Annual Research Conference of the US Bureau of the Census in 1988.

Indeed, as one of the "methodological" sciences, statistics and its intertwined probabilistic foundations have exemplified the shortcomings of traditional, discipline-oriented approaches to indexing.

Bill Kruskal's piece will speak for itself, so I need not say more about the value of a multidisciplinary approach in scholarship. But I do wish to say something about Bill, who is now the Ernest DeWitt Burton Distinguished Service Professor at the University of Chicago.

Born in New York City in 1919, he received a bachelor's degree in mathematics and philosophy from Harvard College, Cambridge, Massachusetts, in 1940. A master's degree in mathematics, from Harvard University, followed the next year. He received a PhD in mathematical statistics from Columbia University, New York, in 1955. His association with the University of Chicago began in 1950, when he joined the Department of Statistics as an instructor. He has been there ever since, serving as chairman of the statistics department from 1966 to 1973 and as dean of the Division of the Social Sciences, most recently from 1980 to 1984. He has been editor and associate editor of the *Annals of Mathematical Statistics* and editor of the *International Encyclopedia of Statistics*. In 1982 he also served as president of the American Statistical Association.

Kruskal's most-cited paper appeared in the *Journal of the American Statistical Association* in 1952. "Use of ranks in one-cri-

terion variance analysis,"⁴ coauthored with W. Allen Wallis, has been cited approximately 650 times. The authors provided a *Citation Classic*[®] commentary on this paper in 1987.⁵ Also highly cited is a 1954 paper, again from the *Journal of the American Statistical Association*, coauthored with Leo A. Goodman: "Measures of association for cross classifications."⁶ This paper has also been cited in over 600 publications and was the subject of a 1979 *Citation Classic* commentary.⁷ Kruskal and Goodman continued their work on this topic; the next two parts to this paper appeared in 1959 and 1963 and have each been cited over 100 times.^{8,9} A fourth installment was published in 1972.¹⁰

In the following speech, Bill discusses cross-disciplinary research in the Census

Bureau, in other federal agencies, and at universities. He examines the considerable advantages, as well as the potential drawbacks, of such activities. As he notes, the joint efforts of statisticians and members of other disciplines have produced advances that, while not yet earthshaking, are quite real. Continued efforts in this direction, whatever the disciplines involved, are bound to bring wide-ranging benefits.

* * * * *

My thanks to Christopher King for his help in the preparation of this essay.

©1989 ISI

REFERENCES

1. Snow C P. *The two cultures and the scientific revolution*. New York: Cambridge University Press, 1959. 58 p.
2. Kruskal W H. The n cultures. *Bureau of the Census. Fourth Annual Research Conference proceedings, March 20-23, 1988*. Washington, DC: US Department of Commerce, 1988. p. 231-7.
3. Tukey J W. A citation index for statistics and probability. *Bull. Int. Stat. Inst.* 40:747-56, 1963.
4. Kruskal W H & Wallis W A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assn.* 47:583-621, 1952.
5. -----, Citation Classic. Commentary on *J. Amer. Statist. Assn.* 47:583-621, 1952. *Current Contents/Arts & Humanities* 9(40):20, 5 October 1987 and *CC/Social & Behavioral Sciences* 19(40):20, 5 October 1987.
6. Goodman L A & Kruskal W H. Measures of association for cross classifications. *J. Amer. Statist. Assn.* 49:732-64, 1954.
7. -----, Citation Classic. Commentary on *J. Amer. Statist. Assn.* 49:732-64, 1954. (Smelser N J, comp.) *Contemporary classics in the social and behavioral sciences*. Philadelphia: ISI Press, 1987. p. 35. (Reprinted from: *Current Contents/Social & Behavioral Sciences* 11(26):14, 25 June 1979.)
8. -----, Measures of association for cross classifications. 2. Further discussion and references. *J. Amer. Statist. Assn.* 54:123-63, 1959.
9. -----, Measures of association for cross classifications. 3. Approximate sampling theory. *J. Amer. Statist. Assn.* 58:310-64, 1963.
10. -----, Measures of association for cross classifications. 4. Simplification of asymptotic variances. *J. Amer. Statist. Assn.* 67:415-21, 1972.

The n Cultures

William H. Kruskal
Ernest DeWitt Burton Distinguished Service Professor
University of Chicago

William H. Kruskal discusses the implications of cross-disciplinary research. Reviewing instances of such activity at the US Bureau of the Census, he examines research combining statistics with other disciplines, such as geography, psychology, and anthropology. He also mentions interdisciplinary research at other federal agencies and in universities. Kruskal calls for the continuation of these efforts.

Introduction

It is always a great pleasure for me to participate in an activity of the Bureau of the Census, especially one so full of accomplishment and promise as this conference. That pleasure is reinforced by noting the presence here of so many collegial friends, and by reflecting on the four prior dinner speakers at these conferences: Janet Norwood, Sidney L. Jones, Vincent Barabba, and Martin Wilk.

My primary theme is the importance of cross-disciplinary research in science generally, and in particular at the Bureau of the Census. I shall touch on past and present cross-disciplinary research at Census, especially with disciplines of the social sciences, and I shall encourage still wider cross-disciplinary activity in the future.

Cross-disciplinary research at the Bureau is, of course, consistent with its high scientific reputation, a reputation that includes individual credit and responsibility for scientific work, encouragement of its professional staff to engage in general scientific publication and discussion, protection from partisan political passions, and, broadly speaking, an excellent level of integrity and openness.

The two cultures. My springboard is a famous 1956 essay by the British scientist and literary figure, the late C.P. Snow.¹ Snow's essay was titled "The Two Cultures," and its central point was despair at the existing gulf between the scientific and the humanistic cultures that Snow saw about him. Snow himself personified a link between those cultures: he had been trained as a scientist and had done research, but he later shifted to become an important novelist. Snow felt that there were all too few such links; he pointed to physicists whose idea of advanced literature is

Dickens' novels and equally to poets without the least glimmer of scientific method.

The n cultures. Of course it is an oversimplification to consider only *two* cultures. As Snow himself says, "The number 2 is a very dangerous number: that is why the dialectic is a dangerous process.... I was searching for something a little more than a dashing metaphor, a good deal less than a cultural map: and for those purposes the two cultures is about right...."¹ (p. 9-10)

But the main diversity of culture for us is the standard one of disciplinary cultures: statisticians, psychologists, physicists, linguists, and so on, swim in different seas. My major theme is that combining the waters of these seas can be indeed productive.

So I am interested in *n* cultures of intellectual disciplines, where *n* lies perhaps between 15 and 50. Of course one could also extend *n* greatly, first to all individuals, since each of us has a little separate cultural sea. As Samuel Butler said, "There are not more stars in heaven than there are worlds of thought within this our own planet...."² So we immediately have an *n* in the millions or billions. But we need not stop there: each moment of time is distinct. George Steiner writes of a herculean flux, "...we never step twice into the stream of our own consciousness; it alters always.... The first person pronoun is a momentary fiction, a momentary arrest in a stream of constant transformative energies."³ (p. 16)

Let me draw back from these heady levels to a down-to-earth hobby that some of us may share: discovering fiction that deals centrally with statisticians or statistics...thus building one kind of link between Snow's two cultures. Needless to say to this audience, there are indeed few novels with statisticians as protagonists; I don't know

whether that rarity is cause or effect for the rather pallid public view of our profession.

Ford Madox Ford. There are, however, a few relatively supportive novels. One of the most striking to me is Ford Madox Ford's *Some Do Not...*, published in 1924.⁴ Its primary protagonist is a government statistician named Christopher Tietjens, and part of the story deals with two statistical moral problems faced by him. Both problems arose from official pressure on Tietjens to cook the data. He berates himself for giving in the first time, but his colleague, Vincent Macmaster, comforts him by saying that there had been no fraud. Tietjens had simply complied with a request from above that a calculation be repeated with a different assumption about rates. Yet Tietjens is an English country gentleman of the old school, and he feels the stain on his honor. Macmaster, by contrast, is a social and economic arriviste with what author Ford considered an undeveloped moral sense.

The second incident, during World War I, was more serious. The statistics bureau with which both Tietjens and Macmaster are connected is under pressure to publish a misleading analysis that underplays the effects of German bombings. Tietjens refuses, but—purely as a technical exercise—shows Macmaster how it could plausibly be done. Macmaster, the little rat, then does it, thus pleasing his political masters. He earned himself a knighthood and other goodies.

Now I realize that members of this audience are unlikely to find themselves in such ethical dilemmas, but nonetheless it is encouraging to find a first-class novel that takes seriously a statistical problem.

Cultural diversity. Let us return to cultural diversity, which seems to me relevant to the census in at least three different ways. First, there are big problems in counting and measuring because of cultural diversity within our nation. Concepts of household, residence, income, even room, are, as you well know, subject to all kinds of different interpretations.

Second, and more to the present point, a census bureau's professional staff is almost bound to be relatively homogeneous. Its members will be quantitatively sensitive, well-educated, law-abiding, and otherwise straight arrows. It is not easy for straight arrows to establish and operate census procedures in a larger society that often departs from straight-arrowhood.

Third, and most particularly my theme, is the cultural diversity among intellectual disciplines. By

bringing in psychologists, linguists, anthropologists, etc., to leaven the flour of statistics and demography, surely a census bureau can improve the quality of its work and the wider understanding of that work.

Past cross-disciplinary cultures at Census. I recognize that I am preaching to the converted, for the Bureau has long sponsored cross-disciplinary research, for example, geographical research as it relates to statistics. As another example among many, I cite Naomi Rothwell's psychological research toward understanding the question and answer process. I recall some anthropological activity, and there is also, I believe, a distinguished line of economics research at Census. One can hardly forget the Bureau's role in computer science. I also keep in mind joint research with mathematics, for example, the combinatorial mathematics of security in a line started a bit north of here by Ivan P. Fellegi⁵ and continued at the Bureau by Lawrence H. Cox⁶ and others.

Census-NSF-ASA. A fine example of Census encouragement of interdisciplinary research is in the Census-NSF-ASA program to improve government-generated social science data. I've been reading the March 1987 proposal to continue that program, a proposal that is most impressive retrospectively, at present, and for the predictable future.

The Research Fellows under this program, which began in 1976, have come not only from statistics and demography; they have also been economists, sociologists, political scientists, and so on. One heart-warming account of the resulting interdisciplinary work and its results is given by former Research Associate Harvey Schwartz in his statement that forms part of a 1986 evaluation. He presents specific examples that may be followed up via the bibliographies that form part of the evaluation.

May I suggest nonetheless that improvement of the social science data base would be well served if the Research Fellows and Research Associates could be drawn from a still wider range of disciplines than at present. Perhaps the program could be extended to anthropology, philosophy, history, linguistics, etc., to provide a richness of interacting cultures beyond those of statistics, demography, economics, geography, and psychology.

Many of you know that this research fellowship program has been broadened in another way to embrace the Bureau of Labor Statistics and the Center for Education Statistics. I foresee a healthy spin-off of fresh interactions among Federal statistics agencies via these fellowships.

Last year's conference. As part of my preparation for this talk, I went systematically through the proceedings of the prior—the Third—Annual Research Conference. It is a most stimulating and satisfying document, and its organizers are to be warmly appreciated. Of the roughly 50 papers, about 35 are primarily statistical/demographic. Next in frequency among the interacting cultures that I noted was Economics, with six papers. I found two papers in Computer Science and two or three dealing with Psychology. Judging from its program, the present conference has a similar distribution of cultures.

Two 1987 papers deserve special mention: one by the historian Margo Conk on lessons from the past, and one by Los Angeles statistician William Diemer on a micro-analysis of housing data. Diemer writes feelingly of the Rashomon-like variability in attitudes toward housing: the tax assessor, the fire chief, the voting registrar, the statistician, etc., perceive housing with rather different eyes and expect different things from a housing census. So Diemer's paper itself is sensitive to the n cultures of our society.

It is hardly surprising that most papers at the Conference were in the mainstreams of statistics and demography. After all, what should a census bureau be or do? Still, one welcomes continuation of the existing movement toward cultural glasnost, the movement to interact more with the cultures, for example, of sociology, anthropology, medicine, history, even philosophy.

Current cross-disciplinary research activity at the Bureau, encouraged by Director Jack Keane, includes undercount behavioral research, which indeed reaches to sociology, linguistics, and anthropology. That research is housed in the Census' Center for Survey Methods Research, itself in the Statistical Standards and Methodology domain. As I understand the Center's role, it is to bring together statisticians, psychologists, political scientists, and others for research that includes measurement methods and nonsampling errors, sensitive to issues related to

Interviewers: selection, training, variability;
Varying interpretations of basic, but amorphous, concepts like that of household;
Cognitive sources of response fluctuation;
Questionnaire design;

and so on. It is a splendid plan, and I understand that Barbara Bailar is to be especially applauded for her role in its growth. Kudos also to Betsy Martin.

There must be other Census groups engaged in cross-disciplinary activity, for example, exhibiting links between geography and statistics, history and statistics, perhaps law and statistics, and so on. I wish that I knew enough about the Bureau to describe such activities in detail.

Other Federal agencies. I have already mentioned the Bureau of Labor Statistics (BLS) and one could readily continue with, for example, BLS's economic research or its recent conference on the wording of questionnaires. I also mentioned the Center for Education Statistics, and there are other agencies that are vigorously nonparochial. Consider, for example, the National Center for Health Statistics and its National Laboratory for Collaborative Research in Cognition and Survey Measurement, an initiative sparked by Monroe Sirken. This Center studies in a cross-disciplinary way such important survey questions as telescoping, distortions created by the survey instruments (called conditioning), and confidentiality; it is hoped that there will be a healthy influence as well on the growth of cognitive psychology. I hope it will not be intrusive if I add a small gratuitous suggestion: do not stop with psychology and statistics.... Other disciplines might well be brought into the picture. (I recognize that there is a natural limit in terms of length of the Center's name.)

The National Science Foundation (NSF) has also been mentioned, and I recur to it. NSF's director, Erich Bloch, is an enthusiast of crossing disciplines. His recent statement, "Changing for the Next Century," stresses "emphasis on multi-disciplinary research, because knowledge is exploding beyond the boundaries of the traditional disciplines."

Without pausing to ask how new that boundary-crossing explosion may be, I note that NSF's much publicized current plans for Science and Technology Research Centers is suffused with cross-disciplinary terminology and enthusiasm. This new program, when and if it is funded, will complement a number of existing programs, including one on Engineering Centers. Of course I should mention a long tradition of NSF cross-disciplinary research support now under the rubric of Measurement Methods and Data Improvement. A central figure there has been Murray Aborn.

The Mathematical Sciences segment of NSF may become increasingly interested in cross-disciplinary activity, and a committee of the Institute of Mathematical Statistics has been hard at work framing possibilities. Major figures in that committee are Jerome Sacks and Ingram Olkin.

There must be many other Federal cross-disciplinary statistical activities, for example, those in the National Institutes of Health, those in the National Bureau of Standards, and those in the now statistically vigorous General Accounting Office. I expect that there are further examples in the Department of Defense, where we see a complex tradition going back to that wonderful post World War II book *The American Soldier*.

Outside government. There are all sorts of cross-disciplinary activities going on, many of them on university campuses. The survey/cognition crossing is especially promising. Social psychologists Norman Bradburn, Howard Schuman, and their colleagues have been hard at work for years on such boundary topics as survey question wording, question order, training of interviewers, effects of interviewer characteristics, etc.

A series of recent papers by Judith Tanur, Elizabeth Loftus, and Stephen Fienberg has constructively dealt with further aspects of the intersection of cognitive psychology and statistics, for example, the comparison of self-reporting with proxy reporting. First steps have been taken toward reconciling laboratory results on forgetting with survey forgetting problems.

The invention of randomized responses may turn out to be a great step forward. My impression is that we await joint work on randomized response by psychologists and statisticians.

On a different front there has been vigorous activity in the study of statistical graphics as it intersects with perception and other psychological domains. Among the participants have been statisticians William Cleveland (Bell Labs), Ronald Thisted (Chicago), and psychologist Reid Hastie (Northwestern). I am sure there must be others, and I predict important discoveries in this area.

Related to questions of graphical display are those connected with the design and execution of statistical tables; if graphics has been the Cinderella of statistics, then table design has been Cinderella's slavey...usually neglected and disdained. There has, however, been a recent growth of interest in tables, a growth pointing to important perceptual questions and to philosophical questions as well, fundamental questions in particular about classifications. The so-called Language of Data project, led at first by the late James Dolby and now headed by Nancy Clark, has brought together statisticians, book designers, psychologists, and others to examine tables of data...how they are used and how they can be improved. This project in fact started in the Federal government, in

the Department of Energy where Lincoln Moses was Administrator of the Energy Information Agency. He, together with his colleague Charles Smith, saw the importance—the general importance—of better understanding for data tables. Help came from the Systems Development Foundation and from statisticians, including John Tukey, and others. In England, psychologist Patricia Wright has done fascinating empirical work on tables, with emphasis on near-term applicability.

Now I wish that this joint work of statisticians with others had already resulted in dramatic advances akin to Newton's laws, to Darwinian evolution, to manned heavier than air flight, and the like. Major contributions are likely to come, I think; in the meantime I note that advances have been incremental yet real, and that they have clearly had effects on the work of the Bureau of the Census.

Negative aspects of cross-disciplinary research. All is not peaches and cream in any diet, and there certainly are potential drawbacks in cross-disciplinary research. Perhaps the first that comes to mind is that simply putting a psychologist and a statistician in adjacent offices and saying "Go to it" may not work well. An analogy might be to proud parents of a beaming marriageable son and the equally proud parents of a beautiful nubile daughter. Pushing the kids toward each other may have just the opposite effect from what the parents want. On the other hand, I am told that in some cultures arranged marriages are the usual thing, that they work out well, and that they have ardent defenders. Another facet of n cultures.

A second concern that cynics might raise is that third rate scientific talents are attracted to cross-disciplinary research, perhaps because standards are fuzzier. Joint research from a psychologist and statistician, for example, might be approved by the psychologist without understanding the statistical part, and vice versa.

No doubt there are such problems, together with risks of faddishness, but I do not make too much of them provided that we are alert.

Meta-cross-disciplinary research. There has grown up an appreciable literature about cross-disciplinary research. A few minutes with our library's search program, for example, turned up several relevant books, including

Problems in Interdisciplinary Studies. A Netherlands Symposium.⁷

*Managing Interdisciplinary Research. An English Conference.*⁸

The papers in these volumes are full of example and analyses. One point of special interest is whether it is intrinsically harder to evaluate proposals for interdisciplinary research support than for traditional research support within a single discipline. I do not have a crisp opinion.

Another example of celebration and analysis of cross-disciplinary research is a marvelous 1949 article on the education of a scientific generalist by Frederick Mosteller, John Tukey, Charles Winsor, and Hendrick Bode.⁹ The article comes at cross-disciplinary research from the educational and preparation viewpoint, but it is surely relevant to our theme, and just as much so today as almost 40 years ago. Indeed, Mosteller, in his recent Pfizer Colloquium Lecture in 1988, argued for the broadening of statistical education and research, with special reference to questions of public policy.¹⁰

Conclusion. We in statistics should feel special pride in successful cross-disciplinary research with statistical leadership. For example, consider R.A. Fisher and his long period at Rothamsted with agricultural scientists. Out of that came great advances in agriculture, and also wonderful statistical innovations (most especially randomization and analysis of variance) that in turn were applicable broadly.

Jerzy Neyman's seminal confidence interval idea arose, I believe, also during joint research with an agriculturalist in Poland.

Everything is not agriculture. Abraham Wald's invention of sequential analysis methods arose from his connections with U.S. Navy acceptance sampling procedures during World War II. Since then, sequential methods have, I believe, become rather less important in the acceptance sampling

domain but have gained considerable currency in medical experimentation.

Few of us are in the league of Fisher, Neyman, and Wald, yet I think the examples carry a clear lesson. Cross-disciplinary research by statisticians and others together can be highly valuable both for statistics and for the other field, be it psychology, political science, chemistry, geography, history, even literature. (In connection with literature, I learned recently from an obituary of A.N. Kolmogorov that he was deeply interested in the stochastic structure of Russian poetry, especially Pushkin's.)

Grant the above, and one is led back to the theme I sounded at the start of these comments. The Bureau of the Census, one of the great statistical organizations of our day, is to be commended for the degree to which it has supported cross-disciplinary research. It has, however, by no means reached a point of diminishing marginal returns, and I hope that more, and more varied, cross-disciplinary research comes under the Bureau's aegis.

I also hope that some future Ford Madox Ford or C.P. Snow writes a novel with a Census statistician as central character. Just think of the dramatic possibilities inherent in the histories of Morris Hansen, Vincent Barabba, John Keane, Barbara Bailar...and no doubt others.

ACKNOWLEDGEMENT

I thank the National Science Foundation for assistance in writing this paper under grant SES 83-03838. I also thank a number of colleagues who responded helpfully to a draft; in particular, I appreciate suggestions from B. Bailar, V. Barabba, N. Bradburn, C. Bush, T. Clemence, L. Cox, S. Fienberg, B. King, L. Lederman, N. Rothwell, M. Sirken, and J. Tanur. The views expressed here are, of course, not necessarily held by many or any of those kind friends.

REFERENCES

1. Snow C P. *The two cultures and the scientific revolution*. New York: Cambridge University Press, 1959. 58 p. (*The two cultures* was first published in *New Statesman*, 6 October 1956.)
2. Turner F M. *Between science and religion*. New Haven, CT: Yale University Press, 1974. p. 180.
3. Steiner G. Some black holes. *Bull. Amer. Acad. Arts Sci.* 41:12-28, 1987.
4. Ford F M. *Some do not...* London: Duckworth, 1924. 329 p. (Published as part of a tetralogy in 1950 by Knopf, New York.)
5. Fellegi I P. On the question of statistical confidentiality. *J. Amer. Statist. Assn.* 67:7-18, 1972.
6. Cox L H. Suppression methodology and statistical disclosure control. *J. Amer. Statist. Assn.* 75:377-85, 1980.
7. Jurkovich R & Paelinck J H P, eds. *Problems in interdisciplinary studies*. Brookfield, VT: Gower, 1984. 200 p.
8. Epton S R, Payne R L & Pearson A W, eds. *Managing interdisciplinary research*. New York: Wiley, 1983. 200 p.
9. Bode H, Mosteller F, Tukey J & Winsor C. The education of a scientific generalist. *Science* 109:553-8, 1949.
10. Mosteller F. Broadening the scope of statistics and statistical education. *Amer. Statist.* 42:93-9, 1988.