Listening. Learning. Leading.®

Ledyard R Tucker's Affair With Psychometrics: The First 45 Years

Bert F. Green

Ledyard R Tucker's Affair With Psychometrics: The First 45 Years

Bert F. Green

Johns Hopkins University

Presented at a special symposium in honor of L. R Tucker

at the University of Illinois, April 25, 1980

Ledyard R Tucker's Affair With Psychometrics:
The First 45 Years

It is an honor to speak about Ledyard Tucker's long and productive career in quantitative psychology. I speak as one of his students, although most of the current generation of psychometricians are at least indirectly his students, so I speak on behalf of psychology generally. Ledyard Tucker has been a central figure in psychometrics for so long that we wouldn't know what to do without him. We won't have to solve that problem now, since the role of professor emeritus means whatever the incumbent wishes it to mean, and in Tuck's case, it plainly means more work. Still, it is fitting at this milestone to take stock, to review what Tuck has done so far.

Two things of great significance in Tuck's life happened in 1934 and 1935. These happen to most college graduates. They go to work and get married, and so it happened with Tuck. He went to work for Thurstone, well-equipped by his undergraduate training in electrical engineering. Thurstone, as you know, was an engineer by training, so he knew a good psychologist when he saw one. You might say that together, Tucker and Thurstone engineered a lot of quantitative psychology. The second major happening was that he married Leona Escudier. Leona soon learned (probably knew from the start) that she had a rival for Tuck's affection. In fact, Tuck has been having an ***affair*** during his entire married life—an affair with psychometrics. His family long ago accommodated Tuck's love of his work. We are all grateful that they have shared him with us.

Tucker's career has three main epochs centering in three places—Chicago and Thurstone from 1934 to 1944, Princeton and ETS from 1944 to 1960, and Champaign-Urbana, and the University of Illinois from 1960 on. At Chicago he helped to develop the theory of factor analysis, among other things. As Thurstone's assistant, he also took over from Harold Gulliksen the task of organizing the numerical calculations. At that time, many clerks at desk calculators did the extensive calculations needed in applying the new methods to psychological problems. At the College Board and ETS, he was mainly responsible for setting up the statistical procedures for test analysis, test equating, item analysis, and the like.

Harold Gulliksen and Tuck were the main figures in ETS's illustrious psychometric fellowship program, but this was more Harold's responsibility than Tuck's—since Tuck was so busy getting ETS off on the right statistical foot. At Illinois, Tuck turned his major efforts to educating a new generation of psychometricians.

His professional contributions do not factor or cluster so nicely. Indeed, I have had a lot of trouble factoring Tucker. I keep getting this huge first factor that dominates all others, that is clearly identified with factor analysis. One of Tuck's most notable characteristics is the tendency to look at all problems from the perspective of factor analysis. Tests can, of course, be factored. But individual differences in judgment also can be factored, the aptitude-achievement distinction can be factored, change can be factored, and even learning curves can be factored. Once in a while he runs across a problem that cannot be treated by factor analysis. But that's all right; he just changes factor analysis until it fits. He did that at least twice—once to deal with two different test batteries given to the same population, which led to inter-battery factor analysis. Again, he dealt with data on several tests given to several persons under several conditions by inventing three-mode factor analysis, a new and powerful technique that is gradually gaining favor as more people start to understand it.

Ledyard Tucker's most important contribution to factor analysis was also his earliest. Together with Thurstone, he gave us correlated factors. Correlated factors were about as palatable then as three-mode analysis is now, but correlated factors are now widely used. Of course, one reason for the slow acceptance may have been that Tucker got correlated factors wrong! Well, not wrong exactly, but certainly contorted. There are three equivalent ways to denote factor loadings on correlated factors. One can report the correlation between the tests and the factors, nowadays called the factor structure. Or one can report the regression weights for estimating the tests from the factors, now called the factor pattern—the loadings on the so-called primary factors. Or, if one is perverse, one can report for each factor the part-correlation of the test with that part of the factor that cannot be accounted for by the other factors, the so-called reference structure. Geometrically, this is the projection of the test vector on the normal to the hyperplane. This perspective is difficult for many people to understand. It takes clear geometric intuition to understand the reference structure, an intuition Tucker has more of than most of us, as did Thurstone, so perhaps they cannot be faulted for failing to foresee the problems others have. Many of Tuck's students have this geometric sense, too. You can tell by the way their hands make 3D graphs as they talk to each other. In fact, it is probably a selection device. Low spatial students need not apply. The rest of us can be thankful for Harry Harman and lesser lights who had a clearer perception of what sells. What sells is factor pattern and factor structure.

Speaking of spatial intuition and geometry, I was looking out the window as my plane approached Champaign, and I think I know why Tucker likes it here. Central Illinois is Cartesian coordinates on a grand scale. And no more perfect two-dimensional hyperplane exists anywhere.

A second early and profound contribution was the discussion of factor rotation by identifying the zeroes in the factor pattern—that is, the vectors on the hyperplanes. The method Tucker talked about yesterday as LSQHYP, least squares hyperplane fitting, was never published, although it appears in the literature as an abstract. Similar work, called a semi-analytic method of rotation, appeared in 1944; then there was further work on the objective determination of simple structure in 1955. The new method, DAPPFR, direct artificial personal probability factor rotation, has not been published. Tucker says some dumb editor rejected it. I don't remember that, but then I am very good at selective forgetting. In ***Psychometrika***'s defense, I have never heard a worse name for a dandy idea than direct artificial personal probability factor rotation.

About your rejection, Tuck, I can only say that you are in good company. When I first took over ***Psychometrika*** from Chet Harris, I inherited the reject files from the past 10 years, and that file has now accumulated. It contains all the big names in psychometrics—Tucker, Gulliksen, Horst, Harris, Kaiser, Meredith, Linn, Torgerson, Green, Humphreys. I'm not sure about Lord and Joreskog, but I am sure about Green because I had to reject one of my own papers because of bad reviews.

Today, I believe Tucker's general approach to factor rotation is gaining favor. For two decades we have followed the Pied Piper of Berkeley down the expedient road of varimax, but students are again recognizing how insightful Tuck's approach was and are trying valiantly to make it foolproof. It is difficult to make anything foolproof because, as someone said, fools are so clever, but that seems prerequisite to installation in SPSS, BMD, and the rest. I am convinced that this approach of hyperplane definition is the only way to solve the oblique rotation problem, if it ever can be solved.

A third seminal paper published early in Tuck's career is modestly titled, ***Maximum Validity of a Test With Equivalent Items.*** This paper is in fact one of the earliest versions of latent trait theory, or, item response theory (IRT) for mental tests. Workers in IRT today are inclined to reference Birnbaum's work in Novick and Lord when needing historical perspective, but, of course Lord's 1955 monograph, done under Tuck's direction, precedes Birnbaum, and Tuck's 1946 paper precedes practically everybody. He used normal ogives for item characteristic curves, as Lord did later. He made a mistake here, too—a verbal one. He called the correlation of test score with the underlying ability scale ***validity***, and so have all who followed, causing endless confusion among the unwashed. It is a shame that he wasn't at the University of Illinois at the time, since then he could have asked Cattell for a different word for validity. Maybe some Greek word for half-god and half-man, for example. A few years ago, in a more mundane style, I suggested ***fidelity***, but by then it was too late. In any case, the work led to IRT which is fundamental to computerized adaptive testing and may be our main salvation for equating tests in the coming era of test disclosure.

Although to my knowledge Tuck did no more work on IRT himself, he has certainly been supportive. And as we heard last night, he did lots of unpublished work on test theory and practice, especially in the area of test equating, a procedure that even few psychometricians know much about. Very few know that there are many different methods of equating in common use—imbedded item equating, anchor test equating, whole test equating by ***spiraling***, which means the use of matched samples, and now IRT equating. These are crossed with one-link and two-link equating. It is a great pity—possibly even a tragedy—that more measurement experts do not understand the issues and the problems here, for they are critical to an informed discussion about test disclosure legislation. When not even quantitative psychologists are aware of the issues, we have a problem. Part of that problem is that this material is not much discussed in the professional literature.

A fourth major theme in Tucker's early work that has continued throughout his career is the problem of computation. Unlike some of our psychometric theorists, Tucker has always been profoundly aware that equations don't really take on meaning until actual data are used and real numbers inserted. The final results depend on doing some calculations, and unless careful attention is given to the calculations, a method may be completely impractical. Our notion of how much calculation is practical has changed over the years, but the issue has always been one of Tucker's priorities.

Tuck was one of the first members of the ACM. We graduate students used to kid about the uncertainty of whether it was the Association for Computing Machinery, or the Association of Computing Machines. He has been most instrumental in encouraging psychologists generally, and his students in particular, to make use of computing machinery. Tucker did his major work at ETS in its formative years in obtaining the test statistics, item analyses, and so on from the old IBM punched card sorters, tabulators and summary punches. Just as today Tuck's students learn how to program the CDC Cyber and any other computer within reach, so his early students (me, for example) learned how to program the old machines. We learned how to wire the plug-boards; we learned how to operate the sorter, the tabulator, and the summary punch. I was one of the first to use the 603 calculating punch, which actually multiplied! That was in 1950, when the IBM 650 was still on the drawing board. Anyway, as part of our graduate course in factor analysis, Tuck showed us how to do a factor analysis on the tabulating equipment. Tuck actually programmed the tabulator, sorter, summary punch, and operator as a sort of four-unit processor to execute an elaborate interactive program designed to do a centroid factor analysis, complete with all the matrix multiplications. We had already learned how to make a tabulator calculate a cross-product matrix by a clever procedure known as progressive digiting, I believe. This procedure involved interchanging the tens and units positions, somehow, together with lots of sorting. Although it is now a dim, dark memory, at the time I understood how it worked. Indeed, I took secret pride in having understood so far everything about the machines and about factor analysis. But Tuck's demonstration of factor analysis by card machine was overwhelming and demoralizing. I was completely lost. He and an assistant were working at the sorter, the tabulator, the summary punch; sorting the summary punch cards—with several decks of cards of different colors, holding intermediate results, piled about. It is my vague recollection that along about Step 42, somebody realized that they had skipped Step 37, but that may be unfair. In any case, I remember concluding that the time was not ripe for factor analysis by card machinery.

By example, Tuck tried to infuse in his students an interest in the nuts and bolts of computing. Sixty years ago it was tabulating machinery with punched cards. Today it is writing computer programs to do the computation. One of my earliest memories of Dr. Tucker was during my first year at Princeton. I was being given a tour of ETS, and I saw him with a small screwdriver, working on one of the new IBM answer sheet readers that ETS was hoping to use to replace the roomful of temporary workers who hand-scored the answer sheets. Tuck and an IBM technician were trying to modify one reader to be less sensitive, to guard against cheaters who placed faint marks on all the answer sheet spaces in an attempt to outwit the scorers.

But Tucker never gave up. In the early 50s, long before Joreskog's Fortran program for maximum likelihood factor analysis, Tucker and Lord programmed MIT's Whirlwind computer for that task, in machine language, in fixed point arithmetic, no less! The time was still not ripe, largely because of storage problems. One medium-sized matrix would fill up the core of the early machines. Imagine doing a factor analysis with only 4,000 words of core for both program and data. Largely through the efforts of Tucker and people like him, IBM and the other computer manufacturers began to see the market for large-storage machines.

Tuck's most recent important work with Linn and Koopman on simulated data matrices is a very important methodological contribution. It can also be looked upon as the culmination of a continued 45-year effort to get adequate computing facilities.

Tucker has made considerable contributions to psychological scaling, though here, for some peculiar reason, he seems to have needed help. Since I'm sure he needed no help solving the problems, the help he needed was undoubtedly to get others interested in the problem and to get the results written up. The Tucker-Messick model of individual differences in multidimensional scaling is widely used; or a variant of it is used. Quite commonly, the judgments of individuals are examined by inverse principal components, to look for consistent differences among individuals that would indicate different psychological structures for the stimuli. One can then follow Tucker and Messick and scale the idealized components, or one can cluster the individuals and scale the cluster centroids. Some people prefer to use 1NDSCAL or its close relative, three-mode factor analysis, to determine individual differences in judgments, but I still find the Tucker-Messick approach easiest to explain to users and most productive in practice.

Other joint work in scaling includes the Diederick, Messick, and Tucker least squares solution for successive intervals, the Schoenemann and Tucker maximum likelihood method for successive intervals with unequal stimulus dispersions, and the Gulliksen and Tucker method for obtaining paired comparisons from multiple rank orders. These and others are all immensely practical contributions to the use of scaling methods.

One of Tuck's worst faults is his slowness. He is especially slow at reviewing manuscripts for ***Psychometrika,*** but he is slow about lots of things. There are two main reasons why people are slow. Either they are not doing enough, or they are doing too much. I once had occasion to visit a research worker who was being too slow about completing some research that I needed. I had never met this person and wanted to see what could be done to speed things up. When I walked into his office, I could see immediately what was wrong. Nothing was out of place. There were no piles of computer output in sight. There were no piles of manila folders. I did see some bound computer output neatly tucked away on a shelf. The books were all neatly in place in the bookshelves. The chairs were neatly around the table. There was no dust. The damn place shined. It was easy to see that this was not the office of a busy person. Need I say that this was not Tuck's office? The last time I visited Tuck, I was lucky to find a place to sit down. It was neat, in its own way, I suppose. That is, the piles were neat. But there certainly were piles. There was no dust. Every time a dust particle settled, someone delivered a new set of computer output to add to the stack, so the dust was constantly stirred up. Tuck has doubtless said, "No," many times, but his office is testimony to the fact that he has said, "Yes," too often. It would be nice, Tuck, if you would say, "No," a little more often— to somebody else, of course.

Tucker had a long and very fruitful association with ETS, first as head of statistical analysis, and later as a research associate, and now as consultant. He set up many of the statistical procedures that are still in place for item analysis and test equating, as well as the more mundane aspects of test scoring and score reporting. He was, and still is, involved in solving many of the major problems of the testing industry.

ETS has always been very confident of its ability to deal with statistical and measurement problems. One reason is that the various workers always knew that if they ever met a problem they couldn't solve, they could always ask Tuck. There is always a danger in asking Tuck a question. If the problem is complex, there is likely to be a long silence. When this happens to you, you might wonder if he is really working on your problem or daydreaming about Hawaii. Actually, of course, the extensive contortions of his face and head make it clear that Hawaii is not on his mind. He is working. When the answer comes, it is always useful, sometimes interesting, and frequently novel. Occasionally it is not an answer to the question you asked. When in exasperation, you point that out to Tuck, he patiently explains that he knows that, but that he answered the question you should have asked.

Another problem with asking Tuck a question is that you may get more than you bargained for. He may offer a whole new way of looking at the problem. This is very creative, in general, but sometimes it implies a whole new research program.

Tuck is not one's image of a clinical psychologist, but he is remarkably observant. As an example, early in his days at the College Board, the board of trustees decided to visit the Princeton operation and meet the people who did the work. This raised the anxiety level of the staff, mainly from a fear that they wouldn't recognize who was who, and it just wouldn't do to call Irving Lorge, "Professor Thurstone." So Tucker sat down with them and gave them a thorough and remarkably discriminating description of each of the prospective visitors. My informant particularly remembers that Thurstone was the one with the big nose. So you might say, that among Tuck's other talents, he is a skilled nosologist.

Tuck is widely liked. To most of the people who know and work with him, he is a great guy. Of course, this is not the ***universal*** opinion. He has enemies and detractors. Only namby-pambies have no enemies, and Tuck is no namby-pamby. He is a great gentleman, but he is not backward about expressing his opinions. Still, Tuck focuses on the issues—his criticisms are not ad hominem; and even his enemies respect him.

Tuck, quantitative psychology has been greatly advanced by your work, published and unpublished, and also by the contributions of your many students. We have all learned a great deal from you in the past. As you now embark on the fourth major epoch of your career, professor emeritus, we confidently expect to continue to profit from association with you and we wish you well on this new venture.