Memories of Lee J. Cronbach

Lee J. Cronbach, a pioneer in the fields of measurement, evaluation, and educational psychology, passed away on October 2, 2001.

This has been a season of loss for us in America, one that has been quickly followed by resolve and focus. We have agreed as a nation to keep doing our work. And so we will. Instead of dealing with an emerging policy or methodological issue at this time, we direct our attention to the memory of Lee J. Cronbach (LJC), our colleague and mentor. For over fifty years, with relentless intellect, wisdom, and energy, LJC fundamentally formed both thought and procedures in educational research methods, measurement, evaluation, and educational psychology. He served as an advisor, colleague and reviewer of CRESST programs and projects. Here are some personal statements from a very few of his associates.

Noreen Webb, UCLA, Colleague and Former Graduate Student

Many people are familiar with Lee J. Cronbach's illustrious career and his many valuable contributions to the social sciences. Fewer may know how seriously he took his job as mentor. He challenged and stretched his students, expecting the same level of serious thought and careful work he demanded of himself. And he always taught by example. Never do research from behind a desk; roll up
Richard Shavelson, Stanford University, Faculty Colleague and Former Graduate Student

In Europe they talk about one’s doctoral advisor as a “father.” In so many important ways that’s true of Lee. I want to think back about the advice father’s given me over the past 30 some years.

Lee and I didn’t really lock horns at Stanford until the dissertation. I’d given him the dissertation draft, along with Nate Gage and Dick Snow (chair), in May of 1970, as I recall. Nate got the dissertation back to me right away... he penned something like: “Good job... you’re done!” Dick took a bit longer getting back to me with the same message, saying that he’d taken time to briefly edit a few parts and gave a couple of suggestions asking me to explain some things in anticipation of the dissertation defense. Now, uncharacteristically, it took Lee a little longer to get back to me. But no sweat... with Gage and Snow signing off... Lee’s praise might not be effusive, you see, he’s not known for effusive praise... Lee got back to me. And I quote: “In some ways I’m the ideal reader for you and in some ways the worst possible.”

Lee closely followed our personal lives, too, and had something to say on just about every occasion.

Lee closely followed the work of his students years, or even decades, after they completed their degrees. A new report or article might spark a three-page missive, complete with multi-colored diagrams, about alternative analytic methods to pursue, deeper research questions to ask, and previous research or theory to build upon—all of which he modestly termed “free associations.”

Lee closely followed our personal lives, too, and had something to say on just about every occasion. Upon the birth of my twins, he offered advice, “Ask for two quarters of maternity leave,” opinions, “Hope you aren’t feeling impelled to set up a matched-pair experiment,” and stories, “Did I ever tell you about [twins] Johnny and Jimmy? Myrtle McGraw wanted to show that acceleration affected development so she persuaded a family to let her take Johnny to her lab for regular physical and mental practice. Parents were concerned that Jimmy would be at a disadvantage so without McGraw’s knowledge they set up their own acceleration practice for him at home. The part I remember is that they had Jimmy jumping off the refrigerator!”

Lee’s mentoring touched so many lives, colleagues as well as students. What a gift to us all.
State Gains Require New Thinking on Validity

Anne Lewis

With unprecedented speed, high-stakes testing has become embedded in state accountability designs and probably will soon be extended through national policies. About 40 states use test scores to hold schools accountable. And as of this writing, both the U.S. House and Senate have approved bills reauthorizing the Elementary and Secondary Education Act that will require annual student testing in reading and math in grades K-8 and hold schools accountable for annual gains in scores on those tests.

But at both state and national levels, a major point has focused on the validity of gains on most state tests. Tenth-grade scores on the Massachusetts Comprehensive Assessment System, for example, rose so sharply in the last year that the public questioned whether the amount of learning has increased, or has narrowed to test-related skills. These rapid gains may stem from awareness by students that the state test results now affect their graduation, but that alone does not indicate whether the increases represent corresponding improvements in student learning. However, the pattern of large increases in scores on state assessments is so widespread that policymakers, educators, and the public are voicing their concerns about the validity of the increases. Further, these gains are only occasionally supported by results from the National Assessment of Educational Progress.

Methods for evaluating gains have not kept pace with the growth of test-based accountability, according to Dan Koretz of the Harvard Graduate School of Education and a long-time CRESST researcher. In a new CRESST report, Koretz and others discuss improved methods for validating gains on all high-stakes tests. The need is extremely important given the expanded practice of providing financial rewards to schools and teachers based on test score increases, and alternately of placing failing schools on probation with the potential for reconstitution.

In this new framework, Koretz et al. identify at least seven types of test preparation that effect gains. Three of these can produce unambiguous, meaningful gains in scores: teaching more, working harder, and working more effectively. Three others, reallocating resources/time to areas to be tested, alignment of instruction to testing, and coaching for tests, can produce either meaningful gains or score inflation, while the final method—cheating—can produce only score inflation. All of these can come into play at the same time.

Koretz also notes that “neither ‘tests worth teaching to’ nor ‘tests aligned with standards’ are sufficient protection against score inflation.” We have instances of state assessments that are thoughtfully developed, aligned, and implemented, yet we observe questionable increases on all of them. The major challenge is separating the parts of the gains that are caused by desired factors versus undesired factors in order to understand what the gains imply about improved student learning.

Another sticky “gains” issue is the lack of stability in individual school gains. For example, in reviewing data from the Colorado Student Assessment Program (CSAP) Robert Linn and Carolyn Haug found that school yearly progress varied significantly depending on the starting point.

(see State Gains, page 7)
Well, I had joined the Stanford University School of Education faculty in summer 1970—the summer of battle with Lee—as an acting assistant professor for 3 years ... in part to teach Lee's famous course on test theory while he was on leave. In my third year at the School, 1972-73, one day Lee rushed into my office with his left arm extended, fingers clutching no less than about 10 different colored transparency pens. As I recall, Lee was working on presenting something from his G theory book, from the covariance chapter, and asking for advice. (Advice? From me? For Lee? A TRAP! A TEST!!)

One last memory of those early days ... Lee called me into his office one day and demanded an explanation for why I was consulting for a test preparation outfit for the Law School Admissions Test. In a fatherly way, upset with his son and making it clear that I'd gone astray, he recalled for me that at the University of Chicago, where he'd been, the faculty did not consult unless the consulting would prove to be an important and unique learning experience. Teaching kids to take the LSAT, in his estimation, wasn't such a case. I explained that acting assistant professors didn't earn much money ($14,000 a year) and Patti and I had a new baby and the money actually helped. Lee thought a moment. He admitted to remembering the days when he and Helen struggled raising a family and gave me the following solution to my problem:

"Why don't you do what I did. I wrote Essentials of Psychological Testing to solve the income problem."

And now, it's the 50th anniversary of Lee's Coefficient Alpha paper. And now Lee's working on a paper to set the record straight ... and I have the privilege of assisting him. And so, not only can I honor my dissertation father ... He has honored me once again.
Eva L. Baker, UCLA Colleague

Far less connected to Lee was I; one of my mentors, Arthur A. Lumsdaine, and Lee came to intellectual blows over analytic techniques to compensate for the lack of true experiments. Nonetheless, I asked—or more accurately, asked Bob Linn to ask—Lee to serve as a “red team member” for an early CRESST institutional proposal, our practice being to seek independent review of our proposals before their formal submission. I actually thought the 18th draft was pretty well done, and expected only a few quibbles. How wrong I was! Lee took our proposal apart, concept by concept, line by line, except for a few projects that he thought were beyond repair. We had allocated one week following the review to revise the document. Bob, Joan Herman and I worked literally night and day to respond and incorporate the majority of Lee’s concerns. Later, when he agreed to work with our Center, I was both excited and terrified. I had also come to know Lee because his commentary was solicited, over the years, on the status of various researchers and measurement experts considered for employment at UCLA. With unblinking accuracy, he was able to identify weaknesses and strengths of some of the very best contributors in the field. He was a man of great generosity as well. Once, he sent me a letter commending an article I had written, evoking both a level of shock and a thrill I didn’t expect to feel after 20 years in the field. More significantly, he also traveled to attend the memorial service in Los Angeles for Leigh Burstein, an act that would have pleased Leigh immensely and endeared Lee Cronbach to me forever.

Ed Haertel, Stanford University, Faculty Colleague

Lee offered me helpful advice on a lot of things over the years. Some good advice that I still have trouble following had to do with saying “no.” He modeled saying “no” quite skillfully. One time I asked him if he would do something, probably serve as a discussant for some symposium, and his reply, as best I remember, was, “You didn’t ask me that in the right way. If you had asked if there was any way I might possibly consider your request, I might have hesitated two or three seconds before saying ‘no’.”

He observed another time that when you get a call asking you to do something, there are lots of immediate reinforcements for saying yes, but none for saying “no.” So, he suggested that I do the following, as I gather he had at one time:

Each time I said “no” to a request, I was to write the thing down I’d said “no” to on an index card and tape it up on the wall near the phone.

Each time I said “no” to a request, I was to write the thing down I’d said “no” to on an index card and tape it up on the wall near the phone. Soon, I’d have a whole wall full of things I would have had to be doing if I’d said “yes,” and seeing the display would make it easier and easier to say “no” the next time.

The most helpful advice on saying “yes” or “no,” though, was more indirect. I was nearing the end of my term on some board that met several times a year, and had been invited to stay on for another term. I asked Lee about it, and after drawing me out a bit, he said, “I think you’ve learned what they have to teach you.”

What better way could there be for a junior faculty member to decide when to move on?

Robert Linn, University of Colorado at Boulder, Colleague

It is no exaggeration to say that Lee Cronbach reshaped the field of educational measurement. The most important topics in the field all depend on his conceptualizations. He stood alone as the field’s most influential thinker. His sphere of influence, of course, was much broader than just measurement.
It included evaluation and broad areas of research in education and psychology. His influence will continue in decades to come, not only from his own work, but also from that of his students—both formal ones and colleagues at CRESST and elsewhere who naturally assumed that role in their interactions with him. His standard of excellence will remain an inspiration, but his sage counsel will be sorely missed.

**Nathaniel L. Gage, Stanford University, Faculty Colleague for More Than 50 Years**

Lee Cronbach and I met early in September 1948, when we arrived at the University of Illinois, Champaign-Urbana, as new faculty members of the Bureau of Educational Research of its College of Education. Lee was an associate professor with tenure and I was an assistant professor without tenure. In 1962, I moved to Stanford, and Lee did the same in 1964. So we were colleagues for an awesome 51 years.

Lee quickly demonstrated his intellectual and professional seniority by receiving the honor of being invited to lecture at the 1948 annual conference of Educational Testing Service. We shared an office for the first few months, during which I got familiar with some of the ways he got so much done.

We worked separately, he on psychometric problems and I on interpersonal perception. He helped me often, whenever I ran into...}

...anomalies, such as higher validity than reliability coefficients. In 1955, we collaborated on a paper, “Conceptual and Methodological Problems in Interpersonal Perception,” which grew substantially out of his solutions to some of the problems on which I’d asked him for help. This was our only joint publication. Many years later, he asked me why I had shied away from further collaboration with him, and I answered truthfully: He was too “smart” for me, so that he had most of the good ideas whenever we did tackle the same problems.

All this suggests too briefly some of Lee’s characteristics, with which many others have had the same experience: intellectual brilliance, willingness to range widely in helping others with their research problems, generosity with his time in helping others.

"We do not store up truths or laws. What social scientists mostly harvest are additional concepts and inquiry skills, along with careful records of events observed.”

That he produced so much impressive work despite often taking time to help others is a reflection of his superabundant intellectual energy.

In sum, Lee provided me and many others with a model of dedication, productivity, rigor, and integrity—one that improved our careers. Let us hope we find his like again.

**Haggai Kuperminc, University of Colorado at Boulder, Colleague**

One can only begin to list Lee Cronbach’s numerous contributions: from groundbreaking work on reliability, generalizability, and validity to a revolution in program evaluation. From an effort to integrate experimental and correlational psychologies to redefining the role of the social scientist. The list goes on. Still, the whole is bigger than the sum of its parts. In the end, Lee’s legacy embodies the 20th-century paradigm shift: from a rationalist, atomistic, and simplistic cause-and-effect view of the world, to the realization of the importance of context, interconnectedness, multiple interacting influences, and plurality of interpretations.

Lee wrote: “We do not store up truths or laws. What social scientists mostly harvest are additional concepts and inquiry skills, along with careful records of events observed.”

Let us hope that we will know how to use Lee’s exceptional harvest of concepts and inquiry skills wisely and responsibly.

The CRESST Line
"Scores with a relatively high percentage of students scoring proficient or advanced in the base year," said Linn, "were likely to have smaller gains than schools with a relatively low percentage of proficient or advanced in the base year." They further noted that "regardless of the starting position, schools that show large gains from year one to year two will generally show a decline in year three, while those that show a decline from year one to year two will generally show a gain in year three."

Recent results on California's Academic Performance Index (API) support Linn and Haug's findings, although the California issue focuses on the validity of yearly target gains. Test scores generally rose across California in 2000/2001; far fewer schools qualified for the API rewards, however, because they failed to reach target gains similar to those that they achieved in 1999/2000. Only 57% of schools met their growth target in 2000/2001 compared to 71% in 1999/2000 (Table 1).

Of particular question are the differential rates between California's elementary, middle, and high schools. Sixty-four percent of elementary schools reached their growth target this year, 51% of middle schools, and 27% of high schools. This differential rate is equally apparent in 1999/2000, causing one to question the methods for defining adequate yearly progress. Setting different growth rate targets for elementary, middle, and high schools, while one alternative, would spark debate about the fairness of setting different changes to other schools may not produce similar gains, nor are such gains likely to be valid increases in student achievement.

Considerable work remains to improve methods for validating gains. It is a high priority for CRESST researchers and states everywhere.

References

In Memoriam

Dr. Sylvia Johnson

Dr. Sylvia T. Johnson, professor of research methodology and statistics in the School of Education at Howard University and recipient of the 1999 CRESST Award for Career Educational Research Contributions, died at her home in Silver Spring, Maryland, on August 10, 2001, from complications of cancer.

Dr. Johnson was editor of the *Journal of Negro Education* and a principal investigator at the Center for Research on the Education of Students Placed at Risk (CRESPAR). Nationally known for her work on measurement and assessment issues, she was a crusader for quality and fair testing for minorities.

In addition to authoring nearly 40 journal articles, Dr. Johnson was active in numerous professional organizations, as well as the National Assessment of Educational Progress and the National Board on Testing and Assessment of the National Research Council.


New CSE Technical Reports

- **Productive Helping in Cooperative Groups**
  Noreen M. Webb, Sydney H. Farivar, and Ann M. Mastergeorge

- **Testing Positive Versus Negative Claims: A Preliminary Investigation of the Role of Cover Story**
  on the Assessment of Experimental Design Skills
  Corinne Zimmerman and Robert Glaser
  CSE Technical Report 554, 2001

- **Assessing Student Representations of Inferential Statistics Problems**
  Nancy C. Lavigne and Robert Glaser
  CSE Technical Report 553, 2001

- **Academic Language and Content Assessment: Measuring the Progress of English Language Learners (ELLs)**
  Robin A. Stevens, Frances A. Butler, and Martha Castellon-Wellington

- **Toward a Framework for Validating Gains Under High-Stakes Conditions**
  Daniel M. Koretz, Daniel F. McCaffrey, and Laura S. Hamilton

See the CRESST Web site, www.cse.ucla.edu, for our complete list of Technical Reports.