

**Roger Shepard** (president 1973-74)

My psychometric interest has always been in methods for discovering and representing the cognitive structures underlying behavioral data. Thus, the methods of nonmetric multidimensional scaling and clustering that I and my Bell Labs colleagues developed during the 1960s and 70s played an important role in my later establishments with my Stanford students in the 1970s and 80s of (1) a “second-order isomorphism” between perception and mental imagery, (2) the geometric structures underlying mental rotation and apparent motion, (3) the geometric structures underlying music perception and cognition, and (4) a “universal” exponential-decay law of generalization—results recognized by the 1995 National Medal of Science and by the 2006 Rumelhart Prize “for outstanding contributions to the formal analysis of human cognition.” In the coming years, I would welcome the development of further methods for the discovery and representation of cognitive structures having forms still more general than those of continuous flat Euclidean,  $L_p$ -norm, or constant-curvature spaces or than those of discrete clusters or tree structures. As examples, I note that some promising advances in these directions have already been made by my former associates J. Douglas Carroll and, most recently, Joshua Tenenbaum and his co-workers.

**J. Douglas Carroll** (president 1975-76)

My 1976 Presidential address “Spatial, non-spatial and hybrid models for scaling” was based on my belief evolution has led to two somewhat but not totally correlated systems for storing and processing information in important stimulus domains; e.g., color. One of these, primarily located in the right brain hemisphere, encodes information in a continuous multidimensional spatial form (e.g., the well-known 3-dimensional color space), while the other, primarily located in the left hemisphere, stores very nearly the same information as discrete “chunks” or clusters (possibly overlapping) of this continuous space, each chunk denoted by a word or other symbol (e.g., color names), by which each can be addressed or referenced. We have all heard the old “saw” that Eskimo languages have many words for different colors of snow that they discriminate quite precisely while we distinguish among them hardly at all. On the other hand, some languages do not have different words for “green” and “blue”, and so tend to confuse the two much more than ours, while others have many words for different shades of blue so that their speakers discriminate among these particular colors, and can refer to them, much better than we can. This seemingly (but not actually) redundant discrete encoding of basically continuous spatial stimulus domains provides the basis, I'd argue, for language. I have, together with various collaborators, devised algorithms for deriving two-way and three-way (individual differences), continuous, discrete and hybrid representations of data based on similarity or other “proximity” data. I look forward with great anticipation to both considerable methodological advances in fitting these measurement models, as well as experimental advances in studying the underlying psychological processes empirically. Ideally this could result in a general model of language in which ordered combinations of words or other discrete symbols connected by set theoretical operations (“and”, “or”, and possibly other string concatenation operators) are used to denote potentially non-contiguous but highly precisely delineated regions (or unions, intersections, or other functions of points and regions) corresponding to analogically encoded “meanings” defined in the continuous perceptual/conceptual space, which would provide the fundamental basis of intra-individual and inter-individual communication. A complementary process could be

postulated in which primitive “meanings” encoded in the continuous spatial representation are translated into a corresponding discrete “verbal” representation, whose validity could be assessed by the extent to which “back-translation” of it accurately maps it back into the original meaning defined by region(s) in the continuous spatial representation. Linguistic processes then could be formulated as interactive alternating mappings between the discrete and continuous representations on either an intra- or inter-individual level. Psychometric methodology in this context would comprise measuring the continuous and discrete psychological variables involved, and using the results of these measurements to test alternative theories of language and communication.

**Duncan Luce** (president 1976-77)

Despite my having been the 1977 president of the Psychometric Society, I am no practicing psychometrician. Rather, I have focused on trying better to understand scientific measurement of easily manipulated attributes and on applying these ideas in psychophysics (e.g., global loudness) and utility theory. The key ideas are having independent measures of an attribute from a concatenation structure and from a conjoint structure that are rendered identical via some sort of “distribution law” between the two structures. Luce (2005) described this perspective a bit and raised some challenges for psychometric theory. So far as I know, this direction has not been pursued despite the value of these methods in other fields. My hope is that ultimately the field will explore more deeply the fundamental laws that must be postulated and tested to lead to principled measurement of such human properties as intelligence and personality traits, presumably resembling physical constants, such as elasticity of a material, rather than as manipulable attributes such as mass or volume. I do not underestimate how difficult this challenge is, but I feel that the field should not underestimate its importance.

Luce, R.D. (2005) Measurement analogies: Comparisons of behavioral and physical measures. *Psychometrika*, 70, 227-251.

**Roderick P. McDonald** (president 1985-86)

In my presidential address (1986) I described a unified and general approach to a large class of psychometric models. The theory was based on a nonlinear model estimated by ML with quasi-Newton minimization methods. Its development required fairly heavy matrix algebra, a symbolic matrix calculus for first and second derivatives of the likelihood function, and linearization of nonlinear models by harmonic analysis. Since then, a new generation of researchers has switched almost unanimously to Bayesian estimation by MCMC. The virtues of this new technology are that it bypasses heavy mathematics and it makes feasible a wider range of specific models. However, it tends to fragment the field and it seems to gloss over some of the classical problems, giving, for example, plausible but in reality unacceptable estimates in models that are not identified. I would like to see a re-integration of the field, with all the advantages of computer-intensive techniques, but also with due recognition of the classical problems and methods. So that psychometrics does not lose its foundation in psychology, careful attention needs to be given to the interpretability of models. Correspondingly, method papers need to include genuine, well-motivated applications.

**Jan de Leeuw** (president 1987-88)

Psychometrics is one of the many subfields of Applied Statistics, with its client population mainly in Education and Psychology. In order to do well Psychometrics has to make sure to strengthen its individual and collective connections with mainstream statistics. Isolationism will eventually lead to inbreeding and deterioration of the gene-pool. And, more importantly, Psychometrics has to strengthen the connections with its client departments. This is relatively unproblematic for Education, where test theory will always be needed, but much more difficult for Psychology. Exciting developments in Psychology are in Bayesian models for cognition and perception, and in data analysis techniques for data from neurological sources, such as fMRI, CAT, and PET. Psychometrics could be helpful there, by shifting its emphasis somewhat from “many observations on few variables” to “few observations on many variables”. Another very active field, which psychometricians could perhaps introduce to Psychology departments, is machine learning and other computer intensive forms of regression. Psychometricians should be aware of the fact that areas such as multidimensional scaling and nonlinear dimensionality reduction have largely been taken over by Computer Science -- they must more aggressively defend the historical accomplishments of Psychometrics. Survival as an independent discipline will depend on maintaining psychometric specializations in Psychology and Education graduate programs. It is important to design and teach a psychometric core curriculum, because without such a core the centrifugal forces will eventually make the discipline disappear. And, of course, suitable sources of funding for client-oriented projects must be found, because that days that psychometric research could be funded from university budgets are clearly over.

**Robert J. Mislevy** (president 1993-94)

When I consider the current state of psychometrics, I see two things that please me and two that raise concerns. The near future of psychometrics seems solid, as the practical importance of our tools and concepts is becoming appreciated for matters of guiding educational and social policy, and in evidence-based decisions in these arenas for individuals. I am excited by advances in modeling and estimation; the long-awaited “Bayesian revolution” seems to have finally arrived. I see less involvement by psychometricians than I would like, however, in simulation-based and game-based learning and assessment. The principled ways we think about modeling persons’ knowledge and skill from uncertain evidence are needed in these new settings. A central reason for this lack, and more troubling, is a growing gap between the language and concepts of psychometrics and the language and concepts of an emerging integration of individual, situative, and social perspectives on cognition—a “sociocognitive” perspective. Working at this level is necessary both to model the actions of subjects in interactive situations, and to collaborate with the communities that are driving these developments.

**Shizuhiko Nishisato** (president 1995-96)

*Psychometrika* used to publish papers mostly from well-known ‘old’ psychometricians, and now many contributors are young and new from a number of countries. This is a welcome change. My concern is with the current unsatisfactory dissemination of newly developed psychometric techniques to the users. First, in the behavioral sciences, most data do not qualify as ratio measurement, yet researchers compute means, variances and correlation, treating their nominal, ordinal or interval data as if they were ratio measurement. A typical example is the use of Likert-type scores, the mean of which, for example, is not ‘optimal,’ for those scores may not reflect the phenomena researchers are supposed to capture with. Secondly, the principle of parsimony has gone too far. Researchers are mostly unaware that (1) multidimensional correlation looks larger than it is if viewed in smaller dimensional space, and (2) the distance between two points in 3-dimensional space, for example, looks closer if viewed in 2-dimensional space. Thirdly, the row space and the column space of a two-way data matrix are not the same: why should we not investigate the relation between, for example, multidimensional anxiety structure and multidimensional subject structure of the subjects-by-anxiety data matrix?

**Fumiko Samejima** (president 1996-97)

My presidential address is entitled, “Departure from Normal Assumptions: A Promise for Future Psychometrics with Substantive Mathematical Modeling.” While normal assumptions are useful working hypotheses, more substantive mathematical modeling is necessary in psychometrics. I still believe in that, and have proposed asymmetric item characteristic function (e.g., logistic positive exponential family of models) on the dichotomous response level, and its expansion to the heterogeneous case of the graded response model, its further expansion to the continuous response model, to give some examples. But not so many reactions have been received. It took twenty to thirty years before my first paper of proposing the general graded response model in 1969 (*Psychometrika* Monograph No. 17) started to be cited by many people, so this time it will take about the same amount of time, and in my life time it may not come. Considering that possibility, I presented a paper at the NCME meeting on “truncated logistic model” and its usefulness, narrowing the focus, and then there were quite a few requests for the paper. I guess I will try this approach, and hope before I die I shall be able to leave something to the younger generations. But I want to say to young people, “Don't depend on software too much that someone else wrote, and study more mathematics while you are still young, so that you will be able to know what you are doing. It will be more fun than you might think.”

**Susan Embretson** (president 1998-99)

My 1999 presidential address represented the merger of research lines that I had been pursuing into a practical outcome. Reliable and valid items could be generated from a computer program to target levels and sources of cognitive complexity! The address concerned the following: 1) predicting the psychometric properties of abstract reasoning items from a cognitive model, 2) demonstrating an item generator based on the cognitive model, 3) developing a new IRT model (the 2PL-Constrained) to predict item responses from the cognitive model and 4) examining how using estimated item parameters impacts ability estimates. Since 1999, many IRT models have been developed to link item responses to substantive features of items, including models for cognitive diagnosis. Research on item generation continues, including a demonstration that GRE items could be generated “on the fly” (Bejar et al., 2003) with adequate psychometric properties. My recent research has found that the psychometric properties of tests based on cognitively based item structures also are predictable. As yet, however, large volume tests have not been produced with item generation methods or cognitively-based item design principles. Hopefully, the methods will be applied soon to solve some current testing problems.

**Wim van der Linden** (president 1999-2000)

Eventually psychometrics will mix with statistics, from which it has already adopted its main tools for modeling, parameter estimation, and model validation. However, its place in statistics will remain unique because of the tradition of modeling of observations for fixed subjects on fixed stimuli it represents---a tradition fed by the S-O-R paradigm in psychology. Mainstream statistics has always had a strong preference for models for identically distributed observations and refrained from modeling that would involve “nuisance parameters” due to differences between subjects or stimuli. But new applications of psychometrics in such areas as biostatistics, medical sciences, survey research, and marketing show the potentially large applicability of psychometric modeling. The true challenge for psychometrics is not to loose its connection to psychology and education. The only way to prevent this is for psychometricians to permanently translate the latest developments in their field into accessible text books and simple computer programs.

**Willem Heiser** (president 2003-04)

The role of psychometrics is that of a mediator, who picks up methodological difficulties that arise in psychological and educational research and tries to solve them by sound mathematical or statistical strategies. On occasion, the mediator might also try to do the reverse, that is, introducing mathematical or statistical strategies that proved useful elsewhere as an innovation for the behavioral sciences. Often, I believe, these attempts fail because of an error of the third kind: giving a correct answer to the wrong problem, or to a problem that does not exist. One example that comes to mind is Bayesian statistics: although it flourishes in statistics, psychologists are sceptical, because they believe that Popper already solved the induction problem—with Fisher’s help—and I cannot blame them. Currently, the greatest challenge for our mediator role is that we no longer have a monopoly, and we experience increasing competition from (a) canned data processing, (b) psychometric experts who do not feel at home in the Psychometric Society, and (c) psychologists who adjust their research paradigms to the statistical tools they happen to know already. Our recurrent research themes (modeling individual differences, dealing with measurement error, and modeling relational and categorical data) are not at stake. But as a community, we should work harder on our mediation skills.

**Roger E. Millsap** (president 2006-07)

I think that to survive as a discipline, psychometrics must be attuned to the needs of researchers in many areas, but particularly in areas that form a natural constituency: education, cognitive science, developmental psychologists who study change, clinicians with assessment interests, and so on. Psychometrics does not do well when it becomes insular, focused solely on refinements of existing methods. That work is necessary, but we also must be listening for new opportunities, new problems to work on. These will emerge from outside psychometrics. A current example is the large data array work from brain imaging research. A further point is that, at least in the United States, we need to attract more young scholars to the field. The demand exceeds the supply at present.