2007 AWARD WINNERS

Distinguished Scientific Contributions

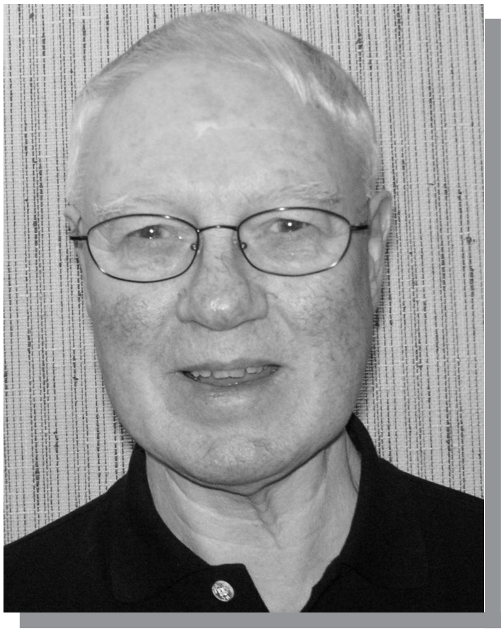
The 2007 recipients of the APA Scientific Contribution Awards were recognized by the 2006 Board of Scientific Affairs and selected by the 2006 Committee on Scientific Awards. Members of the committee were Nora S. Newcombe, PhD (Chair); John F. Disterhoft, PhD; Susan Mineka, PhD; Kevin R. Murphy, PhD; Anne Treisman, PhD; and Timothy D. Wilson, PhD.

Award for Distinguished Scientific Applications of Psychology

The Award for Distinguished Scientific Applications of Psychology is presented to a person who, in the opinion of the Committee on Scientific Awards, has made distinguished theoretical or empirical advances leading to the understanding or amelioration of important practical problems.

1973 Conrad L. Kraft
1974 Gerald S. Lesser and Edward L. Palmer
1975 Nathan H. Azrin
1976 Fred S. Keller
1977 Starke R. Hathaway
1978 Alphonse Chapanis
1979 Joseph Wolpe
1980 Edwin A. Fleishman
1981 Anne Anastasi
1982 Robert M. Gagné
1983 Donald E. Super
1984 Gerald R. Patterson
1985 John Money
1986 Martin T. Orne
1987 Robert Glaser
1988 Leonard Berkowitz
1989 Aaron T. Beck
1990 Wallace E. Lambert
1991 Joseph V. Brady
1992 Charles R. Schuster
1993 Herschel W. Leibowitz
1994 John E. Hunter and Frank L. Schmidt
1995 Ann L. Brown
1996 Ward Edwards
1997 Harold Stevenson
1998/1999 Loren J. Chapman and Jean P. Chapman
2000 David H. Barlow
2001 David T. Lykken
2002 Robert Rosenthal
2003 Stephen J. Ceci and Elizabeth F. Loftus
2004 Edward Taub
2005 Karen A. Matthews
2006 John P. Campbell
2007 Karl G. Jöreskog* Peter M. Bentler*

* This award was shared; this was not an award for collaboration.



Karl G. Jöreskog

Award for Distinguished Scientific Applications of Psychology

Citation

"For his development of models, statistical procedures, and a computer algorithm for structural equation modeling (SEM) that changed the way in which inferences are made from observational data; namely, SEM permits hypotheses derived from theory to be tested. Karl G. Jöreskog defined SEM to include path analysis with latent variables as well as a confirmatory version of factor analysis. His many illustrative examples showed how expected patterns of covariances in SEM can be used to evaluate the fit of models that are derived from prior theory. These examples, along with his models, statistics, and estimation methods, converged to define a new approach for observational (correlational) data."

Biography

Karl G. Jöreskog is professor emeritus at Uppsala University in Uppsala, Sweden. He was born in Åmål, Sweden, in 1935 and did his undergraduate studies at Uppsala University from 1955 to 1957, with a major in mathematics and physics. He received a doctoral degree in statistics at Uppsala University in 1963 with a dissertation entitled "Statistical Estimation in Factor Analysis," a topic suggested to him by Herman Wold. He was a research statistician at Educational Testing Service (Princeton, NJ) and a visiting professor at Princeton University from 1964 to 1971. During these years, he published several articles in *Psychometrika* on the method of maximum likelihood applied to exploratory and confirmatory factor analysis, covariance structure analysis, and multiple group factor analysis. These articles laid the foundation for the LISREL model, which was presented for the first time at the Structural Equation Models in the Social Sciences conference held in Madison, Wisconsin, in November 1970.

In 1971, Jöreskog returned to Sweden to become a professor of statistics at Uppsala University. In 1984, he was appointed a research professor of multivariate statistical analysis, a position he held until his retirement in 2000.

Jöreskog has received three honorary doctoral degrees: from the Faculty of Economics and Statistics at the University of Padua, Padua, Italy, in 1993; from the Norwegian School of Economics, Bergen, Norway, in 1996; and from the Faculty of Psychology at the Friedrich-Schiller-Universität, Jena, Germany, in 2004. He became an honorary professor of finance and economics at Tianjin University, Tianjin, China, in 2006.

Jöreskog is a member of the Swedish Royal Academy of Sciences, a fellow of the American Statistical Association, and an honorary fellow of the Royal Statistical Society. He served as president of the Psychometric Society from 1977 to 1978 and organized the first European Psychometric Society Meeting in Uppsala, Sweden, in 1978.

Jöreskog received the Arnberg Prize from the Swedish Royal Academy of Sciences in 1971; the Ubbo Emmius Medal from the University of Groningen, Groningen, the Netherlands, in 1983; the Educational Testing Service Award for Distinguished Service to Measurement in 1987; the Sells Award from the Society of Multivariate Experimental Psychology in 2000; and the Olaus Rudbeck Medal from Uppsala University in 2005.

Jöreskog has authored several books and numerous journal articles on factor analysis and its extensions and on structural equation modeling. Together with Dag Sörbom, he developed the LISREL computer program.

Selected Bibliography

Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, *32*, 443–482.

Jöreskog, K. G. (1969a). Efficient estimation in image factor analysis. *Psychometrika*, *34*, 51–75.

Jöreskog, K. G. (1969b). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183–202.

Jöreskog, K. G. (1970). Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology*, *23*, 121–145.

Vol. 62, No. 8, 768–769 DOI: 10.1037/0003-066X.62.8.768

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *57*, 409–426.

Jöreskog, K. G. (1973a). Analysis of covariance structures. In P. R. Krishnaiah (Ed.), *Multivariate analysis—III* (pp. 263–285). New York: Academic Press.

Jöreskog, K. G. (1973b). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85–112). New York: Seminar Press.

Jöreskog, K. G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In R. C. Atkinson et al. (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2, pp. 1–56). San Francisco: Freeman.

Jöreskog, K. G. (1977). Factor analysis by least-squares and maximum-likelihood methods. In K. Enslein, A. Ralston, & H. S. Wilf (Eds.), *Statistical methods for digital computers* (pp. 125–153). New York: Wiley.

Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal developmental investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303–351). New York: Academic Press.

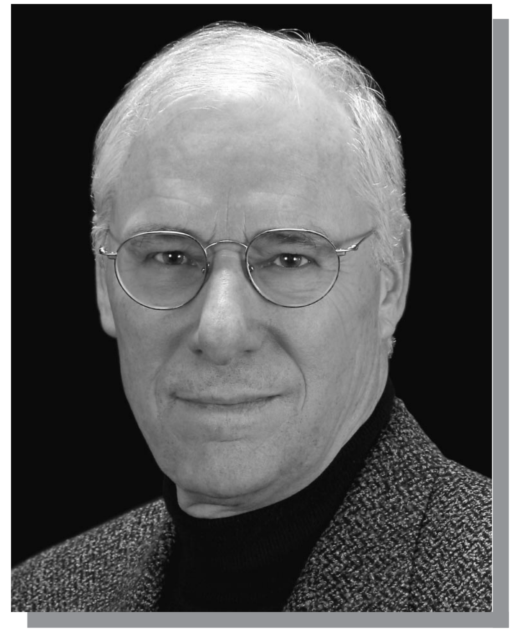
Jöreskog, K. G. (1981). Analysis of covariance structures. *Scandinavian Journal of Statistics*, *8*, 65–92.

Jöreskog, K. G. (1983). Factor analysis as an errors-in-variables model. In H. Wainer & S. Messick (Eds.), *Principals of modern psychological measurement: A festschrift in honor of Frederic M. Lord* (pp. 185–196). Hillsdale, NJ: Erlbaum.

Jöreskog, K. G. (2002). *Structural equation modeling with ordinal variables using LISREL.* Retrieved August 8, 2007, from http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf

Jöreskog, K. G., & Goldberger, A. S. (1972). Factor analysis by generalized least squares. *Psychometrika*, *37*, 243–250.

Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, *36*, 347–387.



Peter M. Bentler

Award for Distinguished Scientific Applications of Psychology

Citation

"For his development of models, statistical procedures, and a computer algorithm for structural equation modeling (SEM) that changed the way in which inferences are made from observational data; namely, SEM permits hypotheses derived from theory to be tested. Peter M. Bentler made SEM available to a broad audience in psychology by providing a model representational system that was conceptually closer to the language of psychological hypotheses. He also developed numerous multivariate statistics and procedures for SEM model testing, model comparison, and model development. His many developments for statistical inference made SEM applicable to the type of observational data that are typical in psychological research."

Biography

Born in Berlin in 1938, Peter M. Blumenfeld lived in Germany during World War II with his brother Claus, mother Ilse, and father Werner, who was twice imprisoned by the Nazis. In 1948, they started a new life in Los Angeles as the Bentlers. Bentler was educated in public schools, including Santa Monica College and the University of California, Los Angeles (UCLA), where he majored in anthropology/sociology and lettered in swimming. After graduation in 1958, he worked as research assistant and then administrator of a small-groups

Vol. 62, No. 8, 769–782 DOI: 10.1037/0003-066X.62.8.769

laboratory at System Development Corporation, an offshoot of RAND. This experience encouraged him to apply to the University of Pittsburgh's program in social psychology. Supported as a research assistant in administrative science (with Professor J. Thompson), he had his first exposure to quantitative methods (R. Glaser, G. Lazovik) and did his master's thesis (T. Wilson, T. Brock) on cognitive dissonance. He became interested in clinical psychology and moved to Stanford University, where he studied depth (K. Colby), behavioral psychology (A. Bandura, L. Krasner), hypnosis (J. Hilgard), and especially personality and individual differences and their assessment (J. Wiggins; also L. Goldberg at Oregon Research Institute). His dissertation was on response variability (1964, with D. Jackson). His graduate publications included an early application of Wolpe's reciprocal inhibition therapy (1962) and five articles on hypnosis. Oddly enough, in view of his over 440 publications, his master's and doctoral theses were never published. After a postdoctoral year in personality assessment at the Educational Testing Service (with S. Messick), Bentler became an assistant professor of psychology at UCLA. Moving through the ranks, he served as chair (1999–2002) and currently is distinguished professor of psychology and statistics.

Bentler maintained an interest in personality and social psychology throughout his career. In his first postdoctorate decade, he was also on the clinical faculty and worked in such areas as sexual behavior assessment, gender constancy and identity, transvestitism, Piaget's conservation, semantic space, and response styles; somewhat later, his focus was on attitudes, love, orgasm, and marital success and failure. In the mid-1970s, he developed an interest in the etiologies, correlates, and consequences of drug use. Integrating his interests in personality and social psychology with research on substance abuse, he designed and implemented a longitudinal study of adolescent development that, with the help of productive collaborators across the years (G. Huba, M. Newcomb, J. Stein, T. Locke), continues today as part of his Center for Collaborative Research on Drug Abuse (http://www.ccrda.psych.ucla.edu), funded by the National Institute on Drug Abuse. As an Institute for Scientific Information highly cited researcher, 40% of his top-cited articles are on drug abuse.

With little formal training in mathematics, Bentler came to quantitative methods through psychology. As he studied psychology in graduate school, he started believing that sophisticated methodologies and quantitative techniques would become critical to its scientific development. For the aspiring psychologist in the early 1960s, Stanford offered many potential directions in career development, as it was becoming a leader in several fields. The growth of mathematical psychology in particular (R. Atkinson, W. Estes, P. Suppes) convinced Bentler that there would be a growing emphasis on quantitative methods in all fields of psychology, requiring him to gain technical proficiency. Encouraged through friendly critiques from correspondents (e.g., H. Kaiser), he augmented his study of statistics and test theory (Q. McNemar) with self-study. As a result, he was able to develop and publish some new results in factor analysis and reliability theory. Although he has worked in several quantitative areas across the years, such as matrix calculus and multidimensional scaling, his most widely recognized quantitative contributions are in the field of latent variable structural equation modeling (SEM).

K. Jöreskog's research of the late 1960s and 1970s was inspirational to Bentler, as his matrix and vector system creatively combined the psychometric factor analytic model with the econometric simultaneous equation system to provide a comprehensive way to specify hypotheses on nonexperimental data. Jöreskog's pioneering LISREL program made it possible for the first time to evaluate such hypotheses in practice. Yet Bentler found that students with average mathematical and statistical backgrounds and those who had trouble with the Greek alphabet had great difficulty learning this methodology; if they were not careful in keypunching their IBM cards, they also had trouble in implementing it. And if their data did not represent samples from multivariate normal populations, the resulting statistics and hence scientific conclusions could be questionable. These issues also directly impacted Bentler's drug abuse research, where skewed and kurtotic distributions were typical. It may be said that his psychometric and statistical work of the past 30 years has been, in good part, directed at understanding and overcoming such limitations.

Although D. Kenny, with his book *Correlation and Causality*, and others initially brought SEM to psychology, Bentler's work was seminal. First, his development of the Bentler–Weeks model provided a much simpler way of explaining and motivating models and permitted logical model specifications, such as the effects of observed variables on others, that were difficult to implement with LISREL. Second, the new Bentler–Bonett fit indices could supplement test statistics to more appropriately evaluate models when sample size is large and power is excessive. Third, his new EQS program allowed researchers to do SEM without needing to know matrix algebra or Greek characters and, later, first permitted the graphical specification of models. Fourth, his early acceptance and extension of the brilliant ideas of M. Browne on distribution-free theory served to assure that SEM would be available for real data that often did not fit the previous restrictive paradigm. Fifth, he and his research group provided many examples of SEM applied to real psychological issues, including drug abuse. And finally, his expository articles brought the ideas of causal modeling to a large audience.

Since then, Bentler has made contributions to virtually all aspects of SEM, especially the associated statistical theory and its performance in practice. His personal contributions include a fundamental definition of latent variable models, models for higher order moments, the comparative fit index, a mixture chi-square test, a noniterative estimator, and new methods in reliability theory. However, most of his publications involve a remarkable set of students and colleagues from around the world. In addition to D. Bonett and D. Weeks, these include M. Berkane, elliptical theory generalization of multivariate normality; W. Chan, models for additive and ordinal ipsative data; C.-P. Chou, model modification and parameter change and Akaike's information criterion; T. Dijkstra, general distribution theory for arbitrary structural models; K. Hayashi, models and methods in factor analysis; L. Hu, simulations on the actual performance of statistics and fit indices; M. Jamshidian, missing data methodology and constrained estimation; Y. Kano, heterogeneous kurtosis theory and pseudo–maximum likelihood theory; K. Kim, homogeneity of means and covariances in missing data; S.-Y. Lee, multimode models, constraint theory, and polychoric–polyserial methods; J. Liang, multilevel models and methods; A. Mooijaart, latent quadratics and interactions and parsimony and precision; W.-Y. Poon, polytomous and interval data models; A. Satorra, scaled and adjusted chi-square statistics and asymptotic robustness; J. Tanaka, limitations of the expectation-maximization algorithm; M.-L. Tang, methods for truncated data and missing data; L.-J. Weng, SEM with dependent observations and multiple population theory; and Y.-F. Yung, bootstrapping methodology and effects of added information.

In the last decade, Bentler worked extensively with K.-H. Yuan, whose many original ideas on finite sample distribution-free and residual-based tests, general multivariate distributions, case-robust SEM, *F* tests, robust missing data methods, distribution-free multilevel tests, mean comparisons, and so on have been unparalleled in the recent history of statistical psychometrics. Bentler is also grateful to have worked and published with students E. Freeman, M. Gold, L. Harlow, G. Hellemann, L. Li, V. Savalei, J. Ullman, and J. Xie; colleagues L. Aiken, P. Dudgeon, K.-T. Fang, F. Faulbaum, T. Raykov, and A. Woodward and many other coauthors. He has been inspired by visiting scholars (M. Browne, B. Byrne, T. Hoshino, A. Klein, G. Marcoulides, E. Meijer, B. Muthén, M. Watanabe, J. Werner, S. West, and K. Yamaguchi); his colleague of 23 years, E. Wu, whose genius helped to create the EQS program; D. Sookne, who cheerfully programs new statistical ideas; supporters at UCLA, especially A. Comrey; and many other students and friends. Certainly not least, his first wife Thea, daughters Katarina and Tania, son-in-law Glen, grandson Jenson, and, for the last 19 years, wife and colleague Elizabeth have provided needed grounding and support.

Bentler is currently president of the Western Psychological Association. Previously, he was president of the Society of Multivariate Experimental Psychology (SMEP); the Psychometric Society; and the American Psychological Association's Division 5, Evaluation, Measurement and Statistics. In 1996, he received Division 5's Distinguished Scientific Contributions Award and, in 2005, SMEP's Sells Award for Outstanding Career Contributions to Multivariate Experimental Psychology. He has received National Institute on Drug Abuse research scientist support for 30 years.

Selected Bibliography

Barnes, G. E., Murray, R. P., Patton, D., Bentler, P. M., & Anderson, R. A. (2000). *The addiction-prone personality.* New York: Kluwer Academic/Plenum.

Bentler, P. M. (1968). Alpha-maximized factor analysis (alphamax): Its relation to alpha and canonical factor analysis. *Psychometrika*, *33*, 335–345.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P. M. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In S.-Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 1–19). Amsterdam: North-Holland.

Bentler, P. M. (in press). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Bentler, P. M., & Speckart, G. (1979). Models of attitude–behavior relations. *Psychological Review*, *86*, 452–464.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, *45*, 289–308.

Jamshidian, M., & Bentler, P. M. (1999). ML estimation of mean and covariance structures with missing data using complete data routines. *Journal of Educational and Behavioral Statistics*, *24*, 21–41.

Kim, K. H., & Bentler, P. M. (2002). Tests of homogeneity of means and covariance matrices for multivariate incomplete data. *Psychometrika*, *67*, 609–624.

Lee, S.-Y., Poon, W.-Y., & Bentler, P. M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology*, *48*, 339–358.

Liang, J., & Bentler, P. M. (2004). An EM algorithm for fitting two-level structural equation models. *Psychometrika*, *69*, 101–122.

Newcomb, M. D., & Bentler, P. M. (1988). *Consequences of adolescent drug use: Impact on the lives of young adults*. Beverly Hills, CA: Sage.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Yuan, K.-H., & Bentler, P. M. (1997). Mean and covariance structure analysis: Theoretical and practical improvements. *Journal of the American Statistical Association*, *92*, 767–774.

Yuan, K.-H., & Bentler, P. M. (1998). Structural equation modeling with robust covariances. In A. Raftery (Ed.), *Sociological methodology* (pp. 363–396). Malden, MA: Blackwell.

Yuan, K.-H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M. E. Sobel & M. P. Becker (Eds.), *Sociological methodology 2000* (pp. 165–200). Washington, DC: American Sociological Association.

Yuan, K.-H., & Bentler, P. M. (2003). Eight test statistics for multilevel structural equation models. *Computational Statistics & Data Analysis*, *44*, 89–107.

Yuan, K.-H., & Bentler, P. M. (2006). Mean comparison: Manifest variable versus latent variable. *Psychometrika*, *71*, 139–159.

Yuan, K.-H., & Bentler, P. M. (2007). Structural equation modeling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 297–358). Amsterdam: North-Holland.

Yuan, K.-H., Hayashi, K., & Bentler, P. M. (2007). Normal theory likelihood ratio statistic for mean and covariance structure analysis under alternative hypotheses. *Journal of Multivariate Analysis*, *98*, 1262–1282.

Can Scientifically Useful Hypotheses Be Tested With Correlations?

Peter M. Bentler
University of California, Los Angeles

Historically, interesting psychological theories have been phrased in terms of correlation coefficients, which are standardized covariances, and various statistics derived from them. Methodological practice over the last 40 years, however, has suggested it is necessary to transform such theories into hypotheses on covariances and statistics derived from them. This complication turns out to be unnecessary, because the methodology now exists to test hypotheses on latent structures of correlations directly. Two examples are given. Limitations of correlation structures are also noted.

Keywords: correlation structures, structural equation models, statistical models

Cronbach's (1957) classic article "The Two Disciplines of Scientific Psychology" highlighted the two different methodological traditions in psychology associated with randomized experimentation on the one hand and analysis of individual differences and relationships among variables on the other. He stated, in part:

> Correlational psychology studies only variance among organisms; experimental psychology studies only variance among treatments. A united discipline will study both of these, but it will also be concerned with the otherwise neglected interactions between organismic and treatment variables. Our job is to invent constructs and to form a network of laws which per-

Editor's Note
Peter M. Bentler received the Award for Distinguished Scientific Applications of Psychology. Award winners are invited to deliver an award address at the APA's annual convention. A version of this award address was delivered at the 115th annual meeting, held August 17–20, 2007, in San Francisco, California. Articles based on award addresses are reviewed, but they differ from unsolicited manuscripts in that they are expressions of the winners' reflections on their work and their views of the field.

> mits prediction. . . . In both applied work and general scientific work, psychology requires combined, not parallel, labors from our two historic disciplines. In this common labor, they will almost certainly become one, with a common theory, a common method, and common recommendations for social betterment. (pp. 681–683)

In subsequent years, Cronbach devoted substantial effort, primarily related to education, to actualizing this vision, especially with regard to aptitude–treatment interactions that require both individual difference and experimental contributions (e.g., Cronbach, 1975). Such a search for variables related to treatment effectiveness remains an important continuing enterprise in many areas of psychological science. For example, in a review of patient–treatment interactions or moderators in evidence-based treatment research, Chorpita, Daleiden, and Weisz (2005) noted, "Collectively, these authors articulated what has become a standard refrain for inference about psychotherapy effects. That is, it is important to know *what works for whom and under what conditions*" (p. 7; see also Roth and Fonagy, 2004). Yet it seems that Cronbach's hope for the larger enterprise of creating a common theory and method integrating both experimental and associational research has not progressed much in the last 50 years. Aside from the difficulties in integrating theories across the varied subdisciplines of psychology, perhaps the field also is simply too broad to permit methodological integration across all areas.

Nonetheless, huge methodological progress has been made on the correlational side of scientific psychology, the topic of this article, especially with regard to the ability to specify and evaluate hypotheses in nonexperimental data. This is not to say, of course, that correlations can prove such hypotheses to be correct; rather, methods for potentially ruling out hypotheses because of their inconsistency with data are now available. Scientific hypotheses may be represented as a structural equation model that specifies how various variables, including unmeasured latent variables (see Bollen, 2002), are presumed to influence other variables in a domain (e.g., a factor analysis model); they may represent specifications that some sets of variables are equally correlated or not at all correlated (e.g., the compound symmetry of equal correlations among repeated measures); or they may represent the idea that a set of correlations can be generated by fewer parameters than the number of observed correlations (e.g., a time series model in which all correlations based on an equal time lag are equal). All such theories can be tested with correlational data, assuming, of course, that the design and general methodology are up to the task.[1] Progress on something so vague and without randomized experimentation is quite remarkable, because a number of psychological issues on humans may remain impervious to experimental study. Illustrative phenomena that cannot be studied by randomly assigning subjects to conditions and following them for years to evaluate outcomes of interest might include the consequences of naturally occurring behaviors such as long-lasting diets low in vegetables, years of heavy cocaine use, or long-term use of performance-enhancing steroids (according to Dr. L. Goldberg, "No one is going to study anabolic steroid use when it's given at 100 times the normal dose"; Stenson, 2007, p. F13). Data on such topics are inevitably correlational. Of course, this means that possible confounding by omitted variables can never be ruled out, as would be possible with randomized treatment research. But the consistency of correlational data with well-formulated hypotheses certainly can now be evaluated. However, the most direct way of doing this—analysis of correlations—does not seem to be popular.

An Old Correlation Matrix and Its Analysis

To illustrate some ideas on the analysis of correlations and to set up the key focus of this article, consider the correlational data presented in Table 1, given almost 50 years ago by Campbell and Fiske (1959). These correlations represent relations among measures of five personality traits (A = assertive, B = cheerful, C = serious, D = unshakable poise, E = broad interests) assessed by each of three methods (1 = staff ratings, 2 = teammate ratings, 3 = self ratings). In their pioneering article, Campbell and Fiske proposed several ways of evaluating such multitrait–multimethod (MTMM) correlation matrices to determine the extent to which they provide evidence of convergent validity of trait measurement across varying measurement procedures and discriminant validity to differentiate possibly related traits from each other. They proposed, for example, that correlations in

> the validity diagonal [*nb* monotrait heteromethod values] should be significantly different from zero and sufficiently large to encourage further examination of validity. This requirement is evidence of convergent validity. Second, a validity diagonal value should be higher than the values lying in its column and row in the heterotrait-heteromethod triangles. (p. 82)

Or

> In the multitrait-multimethod matrix, the presence of method variance is indicated by the difference in level of correlation between the parallel values of the monomethod block and the heteromethod blocks, assuming comparable reliabilities among all tests. (p. 85)

Clearly, this kind of cross-method validation and cross-trait analysis represented an important advance, because it clarified how one might differentiate traits from each other and

[1] This is not a trivial matter, because assumptions, biases, and analytic procedures always remain a concern (e.g., Edwards & Lambert, 2007; Maxwell & Cole, 2007; Yuan & Bentler, 2007).

Table 1
Assessment Study (N = 124)

	A1	B1	C1	D1	E1	A2	B2	C2	D2	E2	A3	B3	C3	D3	E3
A1	—														
B1	.37	—													
C1	−.24	−.14	—												
D1	.25	.46	.08	—											
E1	.35	.19	.09	.31	—										
A2	.71	.35	−.18	.26	.41	—									
B2	.39	.53	−.15	.38	.29	.37	—								
C2	−.27	−.31	.43	−.06	.03	−.15	−.19	—							
D2	.03	−.05	.03	.20	.07	.11	.23	.19	—						
E2	.19	.05	.04	.29	.47	.33	.22	.19	.29	—					
A3	.48	.31	−.22	.19	.12	.46	.36	−.15	.12	.23	—				
B3	.17	.42	−.10	.10	−.03	.09	.24	−.25	−.11	−.03	.23	—			
C3	−.04	−.13	.22	−.13	−.05	−.04	−.11	.31	.06	.06	−.05	−.12	—		
D3	.13	.27	−.03	.22	−.04	.10	.15	.00	.14	−.03	.16	.26	.11	—	
E3	.37	.15	−.22	.09	.26	.27	.12	−.07	.05	.35	.21	.15	.17	.31	—

Note. Data are reproduced from "Convergent and Discriminant Validation by the Multitrait–Multimethod Matrix" by D. T. Campbell & D. W. Fiske, 1959, *Psychological Bulletin, 56,* p. 96, Table 12. Traits are A = assertive; B = cheerful; C = serious; D = unshakable poise; E = broad interests. Methods are 1 = staff ratings; 2 = teammate ratings; 3 = self-ratings.

from method effects as exhibited in correlations. But their approach also was vague and conclusions were subject to interpretation. Confirmatory factor analysis (CFA) brought quite an improvement to the analysis of MTMM matrices, as is well known and as I illustrate with these data.

In terms of the psychological constructs that Campbell and Fiske (1959) were trying to measure, it seems natural to think that there should be five content factors for these data, one for each personality trait. Also, in view of the generally positive correlations across traits within each method block, there also might be some method factors such as halo effects or response sets. I assume that there is one such factor for each method, and, for simplicity, that a given method factor influences its associated observed variables with equivalent strength (equal factor loadings). Thus there ought to be eight factors. Because the traits may well be correlated, I allow them to correlate with each other, but I assume that the method factors are uncorrelated. I also specify that these two different sets of factors are uncorrelated with each other. The setup, as well as results, are given in Tables 2 and 3, which in each cell give the parameter estimate first and the standard error estimate immediately after it in parentheses. The correlation structure chi-square test shows that this model is marginally unacceptable.[2]

The variables used in the analysis are given in the leftmost column of Table 2. I have kept the letter–number designation of each variable rather than emphasizing its psychological meaning, which allows me to focus on the structure of the model setup. The factors that were hypothesized are listed in the top row. Trait A is the factor measured by Variables A1, A2, and A3, and only those three factor loadings in its column were estimated; the rest were fixed at zero. The factor loadings range from .51 to .85, indicating that this factor is well measured by each of the three methods. Trait B is the factor measured by Variables B1, B2, and B3, and again only those three factor loadings in its column were estimated. The loadings range from .41 to .82, again quite reasonable in size. Similar interpretations hold for trait factors Trait C, Trait D, and Trait E, although some of the loadings for Trait D are very low. The last three columns give the factor loadings for the three method factors, Method 1, Method 2, and Method 3, as well as the associated standard error estimates. In each case, all of the different variables involved in a given method, and only those, were used as indicators of the corresponding factor, and the loadings were constrained to be equal.[3]

Various substantive conclusions could be made on the basis of Table 2. For example, Trait D (unshakable poise) is not well measured by Methods 2 and 3 (teammate and self-ratings). In general, the highest factor loadings for trait measures are given by Method 1 (staff ratings), whereas the lowest loadings are given by Method 3 (self-ratings). Method 1 is a very weak factor; in fact, its loadings are

[2] Because raw data are not available to evaluate normality and permit use of a more general test, a normal theory test was used, based on my version of Jennrich's (1970; Shapiro & Browne, 1990) test using an iteratively updated weight matrix. EQS (Bentler, 2002–2007) was used for all analyses in this article.

[3] If my focus was substantive rather than methodological, a series of additional models would be reported on the plausibility of the various restrictions imposed in this model. For example, loosening the equal loading restrictions leads to a significantly improved fit.

Table 2
Confirmatory Factor Loadings and Standard Errors Based on Table 1

Variable	Trait A	Trait B	Trait C	Trait D	Trait E	Method 1	Method 2	Method 3
A1	.850 (.047)					.131 (.107)		
B1		.816 (.059)				.131 (.107)		
C1			.573 (.088)			.131 (.107)		
D1				.822 (.130)		.131 (.107)		
E1					.690 (.079)	.131 (.107)		
A2	.807 (.048)						.297 (.048)	
B2		.672 (.065)					.297 (.048)	
C2			.758 (.090)				.297 (.048)	
D2				.237 (.101)			.297 (.048)	
E2					.659 (.076)		.297 (.048)	
A3	.512 (.073)							.340 (.050)
B3		.414 (.084)						.340 (.050)
C3			.395 (.093)					.340 (.050)
D3				.337 (.098)				.340 (.050)
E3					.512 (.085)			.340 (.050)

Note. All blank entries are fixed zeros. Standard errors are in parentheses. Method factor loadings are fixed to equality. Reweighted least squares $\chi^2(77) = 100.1$, $p = .04$.

not larger than twice their standard errors, so this factor could probably be eliminated from the model. The correlations among all the factors are given in Table 3, which verifies that only the trait factors were allowed to correlate. They are moderately correlated (ranging from −.47 to .63), although two correlations are not significant when judged by their standard errors.

An interesting question is whether this model is simply an example of CFA (Jöreskog, 1969), which has had a long history of application to the analysis of MTMM matrices (e.g., Byrne & Goffin, 1993; Corten et al., 2002; Kenny & Kashy, 1992; Lance, Noble, & Scullen, 2002; Marsh & Hocevar, 1983). The answer is yes and no. The Campbell–Fiske criteria (Campbell & Fiske, 1959) are stated in terms of correlation coefficients: The variances of variables, and hence covariances, are irrelevant. Remember that a population covariance is the number given by the triple product $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, where ρ_{ij} is the correlation between variables i and j and σ_i and σ_j are the variables' standard deviations, and sample covariances are defined equivalently, that is, $s_{ij} = r_{ij}s_is_j$ is a sample covariance with sample correlation r_{ij} and standard deviations s_i and s_j. Clearly, correlations are special cases of covariances with variables that have unit variance. The statistical theory of CFA and MTMM matrices has been given for the analysis of *covariance* structures, based on the statistical distribution of s_{ij} and not on the distribution of r_{ij}, which is appropriate for the analysis of correlation structures. The point of this article is that, as in the example above, when theories are specified via correlations, methods for correlation and not covariance structures are most appropriate. To clarify the issue, a short review of the history of CFA and its extensions is relevant.

Table 3
Factor Intercorrelations and Standard Errors for Model of Table 2

Factor	1	2	3	4	5	6	7	8
1. Trait A	—							
2. Trait B	.546 (.088)	—						
3. Trait C	−.390 (.112)	−.466 (.116)	—					
4. Trait D	.354 (.117)	.631 (.121)	−.048 (.133)	—				
5. Trait E	.542 (.095)	.270 (.121)	.042 (.133)	.472 (.127)	—			
6. Method 1	0	0	0	0	0	—		
7. Method 2	0	0	0	0	0	0	—	
8. Method 3	0	0	0	0	0	0	0	—

Note. Zeros are fixed values. Standard errors are in parentheses.

Factor Analysis, Path Analysis, and Their Extensions

As is well-known, the generally positive intercorrelations among tests and measures of school performance led to the momentous idea that a latent general factor of intelligence could explain such correlations (Spearman, 1904). Of course, a one-factor model is very restricted, and Thurstone (1935, 1947), among others, suggested that a model allowing multiple latent factors would be likely to explain a wider range of psychological data. Today, standard factor analysis is called *exploratory factor analysis* (EFA) in that it aims to find the number of factors that might explain the correlations of observed variables in terms of latent factors, allowing the factor loadings to take on any values. Although EFA might be used to analyze the Campbell and Fiske (1959) data, nothing in the procedure assures that separate trait or method factors could possibly be found rather than some conglomeration of trait and method, and hence EFA is not ideal for this task. In a different domain, Wright (1921, 1934) developed a theory of path analysis to reproduce correlations on the basis of parameters associated with hypothesized causal influences of variables on each other, where the causal influences are represented by standardized regression equations and the size of a standardized beta coefficient quantifies the strength of an effect. These methods were subsequently extended by Jöreskog (1969, 1970) in his method of CFA and his more general higher order model, which in turn were extended further in the merging of CFA, path analysis, and econometric simultaneous equation models (Jöreskog, 1977; Wiley, 1973) as well as in related general linear models (Bentler & Weeks, 1980). Some of this wider structural equation modeling (SEM) history is reviewed in Bentler (1986) and Cudeck and MacCallum (2007). Introductions to SEM are given in many sources such as Byrne (2006) and Raykov and Marcoulides (2006), applications are reviewed by MacCallum and Austin (2000), and recent technical developments in the field are thoroughly discussed in Lee (2007).

Actually, the CFA and SEM literature had made a subtle shift of emphasis. Instead of focusing on correlations, as Spearman (1904), Thurstone (1935, 1947), Wright (1921), and later expounders of correlational methods such as Campbell and Fiske (1959) had done, it focused on covariances instead. Unfortunately, the statistical theory of covariance structures can be problematic at best when applied to the analysis of correlation matrices.

Correlations Versus Covariances

It is, of course, possible to pretend that correlations really are covariances, as largely has been done during the past several decades. There are two main issues with this approach. First, the statistical distribution theory used to develop tests and standard errors can be wrong. Second, key features of the model itself can be destroyed.

The key statistical issue is that the transformation that takes a covariance into a correlation, $r_{ij} = s_{ij}/s_i s_j$, is data dependent, that is, it depends on the sample standard deviations that vary from sample to sample. Such data dependency will modify the statistical distribution of these summary association statistics, or, stated differently, the distribution of r_{ij} may not be well approximated by the formulae that describe the distribution of s_{ij} (e.g., Browne, 1984; Steiger & Hakstian, 1982). Also, in CSA, sample variances s_i^2 have their own sampling variability, but the diagonal elements of a correlation matrix are fixed at 1 and do not vary from sample to sample. Thus, correlations can not be treated as covariances without some potential statistical problems.

The effect of a particular model is more complicated and depends on the model as well as the estimation method used. It is possible for a CSA model to remain valid when applied to correlations, but the conditions for this to hold are hard to describe and evaluate in practice (e.g., Cudeck, 1989; Krane & McDonald, 1978; Shapiro & Browne, 1990). To simplify matters, I ignore the impact of estimation method. Suppose the parameters of a model, such as factor loadings and unique variances, are placed in a vector $\boldsymbol{\theta}$, and we consider the effect of multiplying any variable X_i by k_i. Clearly, this changes its standard deviation from $\sigma_i \rightarrow k_i\sigma_i$, and any covariance involving this variable is also multiplied by k_i. If another variable X_j is similarly scaled, $\sigma_{ij} \rightarrow \sigma^*_{ij} = \sigma_{ij}k_ik_j$. Let us call $\sigma_{ij} = \sigma_{ij}(\boldsymbol{\theta})$ a *scale-invariant* model if there exists another set of parameters $\boldsymbol{\theta}^*$ and constants k_i and k_j such that $\sigma_{ij}^* = \sigma_{ij}k_ik_j = \sigma_{ij}(\boldsymbol{\theta}^*)$ (Browne, 1982; Browne & Shapiro, 1991; Dijkstra, 1990). In a scale-invariant model, the inherent structure of the model is retained when variables are rescaled.[4] This is true of the EFA model, where scaling an X_i by k_i simply rescales the i^{th} factor loadings by k_i. Because a correlation is a rescaled covariance, if a model is scale invariant and an appropriate fit function is used, either a covariance or a correlation matrix can be used to estimate parameters, and the model chi-square test statistic will be identical as well as correct. However, standard errors for some parameters can be correct whereas others may be incorrect. When a model is not scale invariant, rescaling the variables can lead to incorrect parameter estimates, standard errors, and chi-square tests. This occurs, for example, when the model has some fixed nonzero factor loadings beyond those needed for identification, as in a growth curve model or as in my model of Table 2, because it has some equality constraints that would be destroyed by arbitrarily changing the scale of the variables. As noted by Cudeck (1989),

[4] The diagonal elements of the so-called reflector matrix associated with a given discrepancy or fit function are zero for a scale-invariant model (Browne & Shapiro, 1991). This useful diagnostic has not been widely implemented.

> By analyzing a correlation matrix [*nb* with covariance structure methods], one may (a) implicitly alter the model being studied, (b) produce a value of the omnibus test statistic that is incorrect, or (c) report standard errors that are quite discrepant from the correct values. (p. 326)

These problems do not occur when correlations are analyzed directly using appropriate correlation-based statistical theory or when CSA is parameterized so as to embed a correlation structure (e.g., Browne & Shapiro, 1986; de Leeuw, 1983; Fouladi, 2000; Lee, 1985; Mels, 2000; Shapiro & Browne, 1990; Steiger, 1980a, 1980b, 2005; Steiger & Hakstian, 1982). Let $\boldsymbol{S}$ and $\boldsymbol{R}$ be the sample covariance and correlation matrices, and let Σ and $\boldsymbol{P}$ be the corresponding population matrices. Estimates of $\boldsymbol{\theta}$ (say, $\hat{\boldsymbol{\theta}}$) generate $\hat{\boldsymbol{\Sigma}}$ and $\hat{\mathbf{P}}$ with entries $\hat{\sigma}_{ij}$ and $\hat{\rho}_{ij}$, respectively. Then structural models with correlations can be handled by methods that can be correct or incorrect or by methods that are always right (assuming their respective assumptions are met):

1. Model $\boldsymbol{R}$ using CSA. The problems summarized above occur.
2. Model $\boldsymbol{S}$ using CSA and obtain $\hat{\boldsymbol{\Sigma}}$. Transform $\hat{\boldsymbol{\Sigma}}$ to $\hat{\mathbf{P}}$ and interpret the standardized solution that generates $\hat{\mathbf{P}}$. This is a typical practice. It is acceptable only with scale-invariant models. Also, standard errors and *z* statistics for the standardized estimates that generate $\hat{\mathbf{P}}$ should be used. They are known (Jamshidian & Bentler, 2000) but not generally available.
3. Model $\boldsymbol{S}$ using CSA with an embedded correlation structure model $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ where only ρ_{ij} depends on $\boldsymbol{\theta}$, that is, $\rho_{ij} = \rho_{ij}(\boldsymbol{\theta})$, and the population standard deviations σ_i are treated as additional nuisance parameters. Although this is always correct, assuring that $\rho_{ii} = 1$ usually requires nonlinear constraints that are not routinely available in SEM programs.
4. Model $\boldsymbol{R}$ directly using correlation structure methods, with $\rho_{ij} = \rho_{ij}(\boldsymbol{\theta})$. Examples in this article use this method. In EQS, this is accomplished with the command *analysis = correlation* instead of *analysis = covariance*. This approach is always correct.

Note in particular that Methods 3 and 4 are correct for any model for which the given correlation structure model is appropriate, whether or not the model is scale invariant. Indeed, most interesting models for correlations will not be scale invariant. To illustrate what may happen with CSA, I analyzed the correlation matrix in Table 1 with the noninvariant model of Table 2 using maximum likelihood (Method 1), yielding $\chi^2(77) = 105.2$, $p = .02$, a slightly worse fit than with Method 4.[5] Then I rescaled the correlation matrix to an arbitrary covariance matrix and again fit the model of Table 2. The result was $\chi^2(77) = 103.5$, illustrating how the CSA test statistic changes with a non–scale-invariant model.

What If Correlation Structure Methods Had Been Available 40 Years Ago?

If statistical correlation structure methods, that is, Method 4, had been available in the decade following Cronbach (1957) and Campbell and Fiske (1959), the field might not have begun to assume that all correlational problems needed to be rephrased in covariance structure terms. Consider, for example, the field of exploratory factor analysis, which was the dominant multivariate methodology available at the time (for a recent summary, see Jennrich, 2007; Yanai & Ichikawa, 2007). Ever since Spearman (1904), differences in the variances of variables seemed unimportant to detecting latent sources of variability, and many new correlational methods for EFA were developed across the years. The minimum residual, or *minres*, method (Comrey, 1962; Harman & Jones, 1966) was especially promising, because it fit the EFA model directly to the correlations by least squares, no longer requiring the old step of estimating communalities. A variant of Method 4 but without a statistical foundation, it seemed to give meaningful solutions.

A statistical correlation structure theory did not exist at the time. Shortly after minres was made available, Jöreskog (1967) developed a practical CSA-based statistical method for factor analysis that allowed a test for the number of factors and standard errors for factor loadings. Although this development was a cause for rejoicing in the field, because it brought a more refined statistical way of thinking about factor analysis, it required accepting the combined package of statistical methodology plus covariance structure analysis. Even today it is not widely known that a statistical theory exists for general correlation structures.

A leader in factor analysis at the time, Harman (1967) was quick to endorse the new combination of EFA by CSA, to the detriment in growth of his own minres method. After the reanalysis of a classic old data set with Jöreskog's (1967) method, Harman concluded, "For twenty years, two factors had been considered adequate, but statistically two factors do *not* adequately account for the observed correlations based on a random sample of 305 girls" (p. 229).

[5] Also, 16 out of 18 standard error estimates for factor loadings by Method 4 were smaller than those produced by Method 2, sometimes substantially; the other two were equal. For example, the A1–Trait A standard error in Table 2 is .047, but it was .082 by Method 2. This occurs because the standard errors for CSA were not corrected for correlation structures. Cudeck (1989, Table 4) reported that his comparable factor loading standard errors all decreased after correction.

Table 4
Correlations and Residuals for Eight Physical Variables (N = 305)

Variable	1	2	3	4	5	6	7	8
1. Height	—	.846	.805	.859	.473	.398	.301	.382
2. Arm span	−.014	—	.881	.826	.376	.326	.277	.415
3. Length of forearm	−.020	.027	—	.801	.380	.319	.237	.345
4. Length of lower leg	.037	−.020	−.011	—	.436	.329	.327	.365
5. Weight	.016	−.025	.007	.008	—	.762	.730	.629
6. Bitrochanteric diameter	.018	−.006	.011	−.027	.008	—	.583	.577
7. Chest girth	−.021	.004	−.014	.029	.012	−.026	—	.539
8. Chest width	−.024	.044	−.002	−.020	−.027	.021	.015	—

Note. Correlations in the upper triangle are reproduced from "Factor Analysis by Minimizing Residuals (Minres)" by H. H. Harman and W. H. Jones, *Psychometrika, 31*, p. 364, Table 1. Residuals from current analysis appear in the lower triangle.

Today it is possible to disentangle EFA from CSA. The statistical theory of correlation structures can be combined with the minres method to form a new EFA method. This was done by Bentler and Savalei (in press) in their analysis of the correlations among eight physical variables, given in the upper right triangle of Table 4 taken from Harman and Jones (1966). A two-factor minres solution was obtained by using a normal theory least squares method on these correlations. The residuals ($r_{ij} - \hat{\rho}_{ij}$) from this two-factor solution are given in the lower left triangle of Table 4. The minres solution for the factor loadings is the one that minimizes the sum of squares of these residuals. The residuals and minimum function values obtained were identical to those given by Harman and Jones. In addition, the new statistical minres method yielded a normal theory test statistic of 77.5, which, when referred to $\chi^2(13)$, shows that the null hypothesis that two factors can explain the correlations is not tenable statistically.[6] Clearly, there is no need to rely on an awkward covariance structure theory for this result. Yet the EFA covariance structure theory helped to set the stage for CFA, SEM, and so on as they developed across the next 40 years.

For completeness, Table 5 gives the factor loadings for the two-factor solution, along with the estimated standard errors. Clearly, Factor 1 is a height factor whereas Factor 2 is a weight factor. Because height and weight are typically correlated in the population, a further approach to understanding these data might be a confirmatory factor model based on correlated height and weight factors. I do not pursue this direction.

Table 5
Exploratory Factor Analysis for Table 4: Loadings and Standard Errors

Variable	Factor 1	Factor 2
Height	.910 (.174)	.097 (.173)
Arm span	.943 (.221)	.016 (.215)
Length of forearm	.906 (.175)	0
Length of lower leg	.896 (.159)	.070 (.169)
Weight	.412 (.114)	.849 (.200)
Bitrochanteric diameter	.340 (.097)	.724 (.132)
Chest girth	.277 (.095)	.711 (.124)
Chest width	.384 (.080)	.587 (.095)

Note. Standard errors in parentheses. Loading of Variable 3 on Factor 2 fixed at zero for identification. Least squares $\chi^2(13) = 77.5$, $p < .001$, for fit of this two-factor model.

Conclusion

Substantially fostered by my colleague Karl Jöreskog's many contributions, the development of general structural equation methods has remarkably extended the ability to specify and evaluate scientific hypotheses on nonexperimental data. Certainly Cronbach's (1957) suggestion, quoted in the introduction, that "our job is to invent constructs and to form a network of laws which permits prediction" (p. 681) has been widely implemented, even if the thousands of published latent variable SEMs do not rise to the level of laws and, perhaps too often, ignore prediction. In this article, I have emphasized that many hypotheses associated with nonexperimental data are naturally correlational rather than covariance based. As illustrated with two examples, statistical methods now exist to deal directly with the type of correlational questions asked by psychologists 40–50 years ago. Correlational questions are still of relevance today (e.g., Browne, 1992; Preacher, 2006; Steiger, 2005).

[6] Yet, as is clear from the size of the residuals in Table 4, the misspecification with two factors is very minor.

Historically, covariance structure methods have provided a wide range of options for assumptions on the distributions of variables, such as normal, elliptical, heterogeneous kurtosis, or arbitrary distributions, and various specialized test statistics such as the Satorra–Bentler (Satorra & Bentler, 1994) and Yuan–Bentler (e.g., Yuan & Bentler, 1997, 1998) corrections have been developed for improved statistical inference (see Yuan & Bentler, 2007, for a review). Recently, an extended and parallel set of correlation structure methods have been developed and made available (Bentler, 2002–2007; Bentler & Savalei, in press). These allow correlation structures to be applied to the range of realistic data likely to be encountered by psychologists.

Although correlation structure methods now can be implemented in a variety of statistically valid ways, I conclude with some words of caution. First, after several decades of study, a lot of knowledge has accumulated about what type of covariance structure methods might work best under what circumstances. For example, it is now known that the asymptotically distribution-free approach (Browne, 1984) does not work well on large models in relatively small samples with nonnormal distributions; the Satorra–Bentler (Satorra & Bentler, 1994) methodology of using normal theory estimators followed by corrections for nonnormality works much better in this circumstance (e.g., Curran, West, & Finch, 1996; Hu, Bentler, & Kano, 1992). The corresponding knowledge base for correlation structures is minimal (see, e.g., Bentler & Savalei, in press; Fouladi, 2000; Mels, 2000; Steiger, 2005). Further simulation work on the behavior of various correlation structure statistics remains to be done, for example, to determine the most reliable method to use under violations of assumptions such as normality or asymptotic sample size or to evaluate the relative power of various statistics.

Finally, although correlation structure methods are important in their own right, there are indeed times when variances matter and covariance structure methods are more appropriate. One situation is when the basic theory relates to raw score variables rather than to standardized variables, such as in test theory models of parallel measurements. A closely related example is when the scale of the variables has an intrinsic meaning and standardization would destroy that meaning, such as in the study of years of schooling on subsequent annual dollars of income. It might be far more meaningful to conclude that one year of additional college leads to so many dollars of increased income than it would be to say that one standard deviation increase in schooling leads to a 0.3 standard deviation increase in income.[7] The study of factorial invariance also usually requires covariance structures, because equality of covariance-based factor loading matrices across groups is typically a minimum requirement to establish pattern invariance (e.g., Meredith & Teresi, 2006). Another example is multilevel structural equation modeling, in which variances are partitioned into within-cluster and between-cluster sources (Bentler, Liang, & Yuan, 2005; Yuan & Bentler, in press). A further example is growth curve modeling, in which both means and covariances are meaningfully modeled to evaluate individual and group trends across time (e.g., Bollen & Curran, 2006; Duncan, Duncan, & Strycker, 2006). A final example is mixture models, in which means and covariances for one or more hidden groups are modeled simultaneously (e.g., Muthén & Lubke, 2005). In such models, the location and scale of the variables is critical, and correlations then ignore too much relevant information and are not to be recommended.[8]

Author's Note

This research was supported in part by Grants DA00017 and DA01070 from the National Institute on Drug Abuse. Victoria Savalei provided insightful suggestions on early drafts of this article.

Conflict of interest notice: I am a copyright holder of EQS, which was used in all analyses in this article, and a co-owner of its distributor.

Correspondence concerning this article should be addressed to Peter M. Bentler, Department of Psychology, University of California, Los Angeles, 4627 Franz Hall, Box 951563, Los Angeles, CA 90095-1563. E-mail: bentler@ucla.edu

[7] Standardization in regression, equivalent to doing regression in a correlational metric, also has been critiqued (Greenland, Schlessman, & Criqui, 1986; King, 1986) and remains a topic of controversy (Gelman, 2006).

[8] This was also the conclusion Cronbach came to over 50 years ago in another classic article in a different methodological field, namely, assessing similarity between profiles (Cronbach & Gleser, 1953).

References

Bentler, P. M. (1986). Structural modeling and *Psychometrika*: An historical perspective on growth and achievements. *Psychometrika*, *51*, 35–51.

Bentler, P. M. (2002–2007). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Bentler, P. M., Liang, J., & Yuan, K.-H. (2005). Some recent advances in two-level structural equation models: Estimation, testing and robustness. In J. Fan & G. Li (Eds.), *Contemporary multivariate analysis and design of experiments* (pp. 99–120). Hackensack, NJ: World Scientific.

Bentler, P. M., & Savalei, V. (in press). Analysis of correlation structures: Current status and open problems. In S. Kolenikov, L. Thombs, & D. Steinley (Eds.), *Recent meth-*

odological developments in social science statistics. New York: Wiley.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, *45*, 289–308.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605–634.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach.* New York: Wiley.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). London: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*, 62–83.

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika*, *57*, 469–497.

Browne, M. W., & Shapiro, A. (1986). The asymptotic covariance matrix of sample correlation coefficients under general conditions. *Linear Algebra and Its Applications*, *82*, 169–176.

Browne, M. W., & Shapiro, A. (1991). Invariance of covariance structures under groups of transformations. *Metrika*, *38*, 345–355.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum.

Byrne, B. M., & Goffin, R. D. (1993). Modeling MTMM data from additive and multiplicative covariance structures: An audit of construct validity. *Multivariate Behavioral Research*, *28*, 67–96.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, *7*, 5–20.

Comrey, A. L. (1962). The minimum residual method of factor analysis. *Psychological Reports*, *11*, 15–18.

Corten, I. W., Saris, W. E., Coenders, G., van der Veld, W., Aalberts, C. E., & Kornelis, C. (2002). Fit of different models for multitrait-multimethod experiments. *Structural Equation Modeling*, *9*, 213–232.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*, 116–127.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, *50*, 456–473.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, *105*, 317–327.

Cudeck, R., & MacCallum, R. C. (Eds.). (2007). *Factor analysis at 100: Historical developments and future directions.* Mahwah, NJ: Erlbaum.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.

de Leeuw, J. (1983). Models and methods for the analysis of correlation coefficients. *Journal of Econometrics*, *22*, 113–137.

Dijkstra, T. K. (1990). Some properties of estimated scale invariant covariance structures. *Psychometrika*, *55*, 327–336.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytic framework using moderated path analysis. *Psychological Methods*, *12*, 1–22.

Fouladi, R. T. (2000). Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate nonnormality. *Structural Equation Modeling*, *7*, 356–410.

Gelman, A. (2006). *Scaling regression inputs by dividing by two standard deviations* [Working paper]. Accessed

May 15, 2007, from the Society for Political Methodology Web site: http://polmeth.wustl.edu/workingpapers.php

Greenland, S., Schlessman, J. J., & Criqui, M. H. (1986). The fallacy of employing standardized regression coefficients and correlations as measures of effect. *American Journal of Epidemiology*, *123*, 203–208.

Harman, H. H. (1967). *Modern factor analysis* (2nd ed.). Chicago: University of Chicago Press.

Harman, H. H., & Jones, W. H. (1966). Factor analysis by minimizing residuals (minres). *Psychometrika*, *31*, 351–368.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, *112*, 351–362.

Jamshidian, M., & Bentler, P. M. (2000). Improved standard errors of standardized parameters in covariance structure models: Implications for construct explication. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment* (pp. 73–94). Dordrecht, the Netherlands: Kluwer Academic.

Jennrich, R. I. (1970). An asymptotic chi-square test for the equality of two correlation matrices. *Journal of the American Statistical Association*, *65*, 904–912.

Jennrich, R. I. (2007). Rotation algorithms: From beginning to end. In S.-Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 45–63). Amsterdam: North-Holland.

Jöreskog, K. G. (1967). Some contributions to maximum likelihood exploratory factor analysis. *Psychometrika*, *32*, 443–482.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183–202.

Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika*, *57*, 239–251.

Jöreskog, K. G. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P. R. Krishnaiah (Ed.), *Applications of statistics* (pp. 265–287). Amsterdam: North-Holland.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait–multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, *112*, 165–172.

King, G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, *30*, 666–687.

Krane, W. R., & McDonald, R. P. (1978). Scale invariance and the factor analysis of correlation matrices. *British Journal of Mathematical and Statistical Psychology*, *31*, 218–228.

Lance, C. E., Noble, C. L., & Scullen, S. E. (2002). A critique of the correlated trait–correlated method and correlated uniqueness models for multitrait–multimethod data. *Psychological Methods*, *7*, 228–244.

Lee, S.-Y. (1985). Analysis of covariance and correlation structures. *Computational Statistics & Data Analysis*, *2*, 279–295.

Lee, S.-Y. (Ed.). (2007). *Handbook of latent variable and related models.* Amsterdam: North-Holland.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*, 201–226.

Marsh, H. W., & Hocevar, D. (1983). Confirmatory factor analysis of multitrait-multimethod matrices. *Journal of Educational Measurement*, *20*, 231–248.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*, 23–44.

Mels, G. (2000). *Statistical methods for correlation structures.* Unpublished doctoral dissertation, University of Port Elizabeth, Port Elizabeth, South Africa.

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, *44*(Suppl. 3), S69–S77.

Muthén, B., & Lubke, G. H. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21–39.

Preacher, K. J. (2006). Testing complex correlational hypotheses with structural equation models. *Structural Equation Modeling*, *13*, 520–543.

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling.* Mahwah, NJ: Erlbaum.

Roth, A., & Fonagy, P. (2004). *What works for whom? A critical review of psychotherapy research* (2nd ed.). New York: Guilford Press.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Shapiro, A., & Browne, M. W. (1990). On the treatment of correlation structures as covariance structures. *Linear Algebra and Its Applications*, *127*, 567–587.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Steiger, J. H. (1980a). Testing pattern hypotheses on correlation matrices: Alternative statistics and some empirical results. *Multivariate Behavioral Research*, *15*, 335–352.

Steiger, J. H. (1980b). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251.

Steiger, J. H. (2005). Comparing correlations: Pattern hypothesis tests between and/or within independent samples. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 377–414). Mahwah, NJ: Erlbaum.

Steiger, J. H., & Hakstian, A. R. (1982). The asymptotic distribution of elements of a correlation matrix: Theory and application. *British Journal of Mathematical and Statistical Psychology*, *35*, 208–215.

Stenson, J. (2007, March 19). Making the case for steroids. *Los Angeles Times*, p. F13.

Thurstone, L. L. (1935). *The vectors of mind.* Chicago: University of Chicago Press.

Thurstone, L. L. (1947). *Multiple factor analysis.* Chicago: University of Chicago Press.

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 69–83). New York: Seminar.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *10*, 557–585.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*, 161–215.

Yanai, H., & Ichikawa, M. (2007). Factor analysis. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 257–296). Amsterdam: North-Holland.

Yuan, K.-H., & Bentler, P. M. (1997). Mean and covariance structure analysis: Theoretical and practical improvements. *Journal of the American Statistical Association*, *92*, 767–774.

Yuan, K.-H., & Bentler, P. M. (1998). Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*, *51*, 289–309.

Yuan, K.-H., & Bentler, P. M. (2007). Structural equation modeling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 297–358). Amsterdam: North-Holland.

Yuan, K.-H., & Bentler, P. M. (in press). Multilevel covariance structure analysis by fitting multiple single-level models. *Sociological Methodology.*