Graduate Training in Statistics, Methodology, and Measurement in Psychology

A Survey of PhD Programs in North America

Leona S. Aiken — *Arizona State University*
Stephen G. West — *Arizona State University*
Lee Sechrest — *University of Arizona*
Raymond R. Reno — *Arizona State University*

With Commentary by

Henry L. Roediger III — *Rice University*
Sandra Scarr — *University of Virginia*
Alan E. Kazdin — *Western Psychiatric Institute*
Steven J. Sherman — *Indiana University*

ABSTRACT: A survey of all PhD programs in psychology in the United States and Canada assessed the extent to which advances in statistics, measurement, and methodology have been incorporated into doctoral training. In all, 84% of the 222 departments responded. The statistical and methodological curriculum has advanced little in 20 years; measurement has experienced a substantial decline. Typical first-year courses serve well only those students who undertake traditional laboratory research. Training in top-ranked schools differs little from that in other schools. New PhDs are judged to be competent to handle traditional techniques, but not newer and often more useful procedures, in their own research. Proposed remedies for these deficiencies include revamping the basic required quantitative and methodological curriculum, culling available training opportunities across campus, and training students in more informal settings, along with providing retraining opportunities for faculty. These strategies also require psychology to attend carefully to the human capital needs that support high-quality quantitative and methodological training and practice.

During the 1960s, psychologists liked to fancy themselves as the leaders among the social sciences in statistical, measurement, and design issues. Although some of this self-perception undoubtedly stemmed from disciplinary chauvinism, it may have been in part true. Psychologists were then the only social scientists to use factorial experimental designs in the laboratory. Psychologists had made major contributions to measurement, notably psychophysics (Stevens, 1951, 1961), scaling (Coombs, 1964; Torgerson, 1958), and classical test theory (Gulliksen, 1950; Lord & Novick, 1968). Also, they had achieved considerable expertise in complex analysis of variance designs (Winer, 1962) and had been instrumental in the development of factor analysis (Harman, 1967). Moreover, many psychology departments had one or more faculty members with considerable expertise in statistics, methodology, and/or measurement who were widely sought after as consultants by other social scientists.

The training of graduate students reflected these advances, apparently providing a strong basis for addressing the research problems of the day. In reconstructing the training of the 1960s, we have been forced to rely on the admittedly imperfect recollections of older faculty members. Nonetheless, there is relatively good consensus that most departments routinely taught analysis of variance and correlation/regression analysis, with a smaller number offering factor analysis. Measurement seemed to be well represented, with many departments offering courses in measurement theory, (classical) test theory, psychophysics, scaling, and attitude measurement. The level of methodological training is harder to estimate, as the nature of this training varied by area. The result of such training, at least in the recollections of faculty members, was that typical students were adequately equipped to address statistical, measurement, and methodological problems in their areas of research, and the best students were very well equipped to tackle virtually any research problem with which they were likely to be confronted.

The past 20 years have witnessed important developments in statistics, methodology, and measurement. In statistics we have witnessed the development of such techniques as structural equation modeling (e.g., Bentler, 1986; Duncan, 1975; Jöreskog & Sörbom, 1979); discrete multivariate analysis, including logit and probit regression models for use with categorical outcome variables and correspondence analysis (e.g., Aldrich & Nelson, 1984; Bishop, Feinberg, & Holland, 1975; Greenacre, 1984); and time series analysis (e.g., Box & Jenkins, 1976; McCleary & Hay, 1980). The area of measurement has

Vol. 45, No. 6, 721–734

seen the development of multidimensional scaling techniques (e.g., Schiffman, Reynolds, & Young, 1981; Shepard, Romney, & Nerlove, 1972), item response theory (e.g., Hambleton & Swaminathan, 1984; Lord, 1980), and generalizability theory (e.g., Cronbach, Gleser, Nanda, & Rajaratnam, 1972). Methodology has seen the identification of new issues and new design solutions in research, particularly as reflected in large-scale social experiments (e.g., Riecken et al., 1974), longitudinal research (e.g., Kessler & Greenberg, 1981; Nesselroade & Baltes, 1979), and quasi-experimentation (e.g., T. D. Cook & Campbell, 1979; Judd & Kenny, 1981). The development of meta-analysis has allowed the quantitative analysis of entire research literatures (e.g., Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Hunter, Schmidt, & Jackson, 1982).

Unfortunately, however, disciplines other than psychology have often been at the forefront of advances in social science methods. Sociologists and econometricians have provided much of the impetus for statistical advances, and educational researchers have provided much of the leadership in the development of modern measurement theory. Psychology has been most visibly the leader in the development of methodology, led largely by the contributions of Campbell, Cook, and their colleagues and students, and in meta-analysis (in collaboration with educational researchers), although it has made important contributions in certain areas of measurement and structural equation modeling.

Substantive developments have also changed the nature of the questions psychologists ask and the settings in which they perform their research. To cite but three examples, the rise of cognitive psychology has led to interests in theoretically proposed structures and processes whose existence can only be inferred from complex patterns of relationships. Also, the rise of life-span developmental perspectives has led to interests in studying processes that take place over long periods of time. And the rise of a variety of applied perspectives, such as health psychology and community psychology, has led to increased interest in studying phenomena in their real-world settings. Research psychologists, even those trained in the core areas of experimental psychology, have frequently taken applied positions in industry and government (A. Howard et al., 1986), changing the nature of the research problems that they confront. In short, these developments have led to the posing of broader and more complex questions, often addressed with research carried out in nonlaboratory contexts in which traditional designs and analyses provide at best nonoptimal and at worst wrong answers.

In light of the potentially important implications of these developments, one would expect the advances to be incorporated into the research methods curriculum. To evaluate the extent to which this has taken place, we surveyed all departments offering the PhD in psychology in the United States and Canada to assess training in statistics, measurement, and methodology. We raised three general questions concerning training:

1. What is the current content of the general statistics, measurement, and research methodology curriculum in doctoral programs[1] in psychology?
2. What do the various subdisciplines within psychology require specifically in their more focused training?
3. With what methods, analyses, and measurement approaches are new PhDs judged to be sufficiently competent so that they can use these techniques in their own research?

Two additional questions were suggested by changes in the demographics of training in psychology in the United States (A. Howard et al., 1986).

4. There has been a substantial increase in the number of clinical and other practitioner-oriented graduate students relative to nonpractitioner-oriented students. Are there substantial differences in the statistics, measurement, and methodology curricula studied by these two types of students?
5. A decreasing percentage of the new PhDs in the United States are granted by "elite" institutions.[2] What are the primary differences in training between elite and other graduate programs?

Conducting the Survey

Procedure and Respondents

Questionnaires were mailed in February 1986 to the chairs of 222 psychology departments or schools identified by the American Psychological Association (APA) as offering the PhD degree. The accompanying letter requested that chairs provide demographic information on the department and that quantitative/methodological faculty complete the remainder of the questionnaire. Reminder postcards were sent to nonrespondents five weeks later. Two months thereafter a second questionnaire was mailed to all remaining nonrespondents. Respondents who returned incomplete questionnaires were interviewed by telephone to gather more complete information. A total of 186 departments returned usable questionnaires, constituting an 84% completion rate.

Preliminary versions of the results of the survey and the commentary were presented in L. S. Aiken and S. G. West (chairs), *Adequacy of Methodological and Quantitative Training: Perspectives of the Disciplines*, symposium presented at the meeting of the American Psychological Association, New York, August 1987.

We thank Lyle Jones, David A. Kenny, Robert L. Linn, and Howard Wainer for their input to the content of the questionnaire. We thank Peter Bentler, Sanford Braver, Thomas D. Cook, and Joseph Hepworth for their comments on an earlier version of this article.

Raymond R. Reno will be at the University of Notre Dame as of January 1991.

Correspondence concerning this article should be addressed to Leona S. Aiken or Stephen G. West, Department of Psychology, Arizona State University, Tempe, Arizona 85287.

[1] In this article, department, school, and program are used interchangeably. Specific subdisciplines (e.g., developmental) are referred to as areas.

[2] In this article, the "elite" institutions are the 24 departments with the highest reputation ratings in the Jones, Lindzey, and Coggeshall (1982) study of graduate training programs in psychology (see also G. Howard, Cole, & Maxwell, 1987).

Questionnaire

The questionnaire collected demographic information about each department as well as its offerings in statistics, measurement, and research methodology. Demographic questions focused on the number of first-year students, the number of full- and part-time faculty, the use of quantitative/methodological course offerings from other departments, and the existence and nature of a PhD-level quantitative area within the department. The existence and duration of a number of standard statistics, measurement, and methods courses in the curriculum of the psychology department and in the university were assessed; and the curricula taken by practitioner-oriented students and nonpractitioner-oriented students were compared. The number of courses in methods and in statistics and measurement required by each substantive area within psychology was compared, and the coverage of the required introductory statistics sequence was assessed. Finally, the respondents estimated the percentage of graduating doctoral students in their departments who could apply basic and advanced procedures in statistics, measurement, and methodology to their own research.

Results of the Survey

Program Demographics

Program demographics are presented in Table 1. Of the 186 responding programs, 78% had a clinical area, a counseling area, or both, among which were 11 free-standing clinical programs; 88% had at least one academic area (experimental, social, developmental, or personality); 40% had an applied area; and 7% had a quantitative area. Including all tenure track and nontenure track faculty, there was a median of 24 faculty per department. The median first-year class size was 15. Finally, using the reputation ratings from the Jones, Lindzey, and Coggeshall (1982) study, no differences in reputation ratings were obtained between responding and nonresponding departments. This result suggests that the present sample did not overrepresent unusually high- or low-quality training programs.

Table 1
Program Demographics

Characteristic	All programs (n = 186)	Elite programs[a] (n = 21)
First-year class size (*Mdn*)	14.7	17.8
Full-time faculty (*Mdn*)	24.2	31.2
Full-time faculty who teach statistics or measurement (*Mdn*)	2.6	2.8
Students regularly take statistics outside department (% yes)	31%	48%
Program offers PhD in quantitative area (% yes)	17%	43%
Faculty teaching statistics who were trained in statistics or research methods (*Mdn*)	0.9	1.6
Faculty teaching statistics exclusively (*Mdn*)	0.5	1.1
Department offers introductory graduate statistic sequence (% yes)	89%	86%
Of departments offering sequence:		
(a) How long is sequence? (% one year in length)	77%	78%
(b) Is course required of all PhD students? (% yes)	93%	89%

[a] Twenty-one (of 24) elite programs responded to the survey.

Quantitative Staff and Resources

Across all departments, a median of 2.6 full-time faculty teach statistics, measurement, or both at the graduate level. Half of the departments have at least one member who exclusively teaches quantitative courses. One sixth of the departments employ part-time or adjunct faculty to teach PhD-level statistics or measurement courses. Almost one third regularly have PhD students take statistics or measurement courses in other departments, predominantly mathematics (67%) and education (51%).

The backgrounds of the faculty teaching graduate statistics and measurement courses were also examined. A median of 0.9 of the 2.6 faculty members teaching quantitative courses was trained primarily in statistics or research methods, whereas the remainder were trained primarily in a substantive area. One third of the programs have no faculty primarily trained in statistics and/or measurement who teach quantitative courses. These programs are no more likely to have students regularly take quantitative courses outside of the department than programs in which one or more of the faculty had received their primary training in a quantitative area (30.4% vs. 31.3%).

Curriculum Offerings

We first asked about the general curriculum in statistics, measurement, and research methodology available in the respondent's department and on campus. Respondents were presented with a series of course topics representing traditional and newer areas in research methods. Table 2 summarizes these results. As can be seen in column 1, most of the old standards of statistics—analysis of variance (ANOVA), multiple regression, and multivariate analysis—are included in the curriculum of a strong majority of the departments. Among the old standards, only factor analysis is a minority offering. A general research methods course is available in the curriculum of 70% of the departments, whereas more specialized methods courses, such as survey research and quasi-experimental design, are offered in only about one third of the departments. Computer applications has become increasingly popular, with over 50% of the programs offering such a course. Measurement courses, however, are offered by fewer than half of the departments. Causal (structural equation) modeling, a statistical development that is enjoying rapidly increasing use across social science disciplines and within many of psychology's substantive areas,

Table 2
The Statistics, Methodology, and Measurement Curriculum of Doctoral Programs in Psychology

Course area	Coverage (%) Partial course[a] All[c]	(Elite)[d]	Full course[b]	None	Available on campus
Old standards of statistics					
Analysis of variance	88%	(86)%	65%	3%	69%
Multiple regression	68	(71)	36	8	77
Multivariate analysis	63	(67)	48	14	72
Factor analysis	36	(33)	20	20	54
Methodology					
Research methods	70	(62)	56	13	58
Evaluation research	36	(24)	29	37	51
Quasi-experimental design	28	(29)	14	23	42
Survey research	15	(14)	10	48	53
Measurement					
Test theory	45	(48)	31	30	46
Test construction	25	(14)	13	40	42
Scaling	21	(43)	16	38	39
Other topics					
Computer applications	53	(57)	41	13	62
Mathematical psychology	24	(67)	22	64	15
Causal modeling	18	(33)	14	45	48
Time series	6	(5)	4	63	52

Note. Values represent the percentage of schools responding affirmatively to the question.
[a] At least half semester or full quarter. [b] Quarter, trimester, or semester. [c] All schools responding to the survery. [d] Percentages for the 21 responding elite schools are in parentheses.

may be used as an indicator of the adoption of new developments in statistics. As yet, this technique has made few inroads into the curriculum.

Another perspective on these results is provided by examining what is *not* represented in the curriculum, *even for a small segment of a course.* A glance at the measurement topics reveals that over one third of the programs currently offer absolutely no training in measurement. With regard to developments in methodology, nearly one fourth of all departments offer no coverage of quasi-experimental design.

We recognize that many departments, particularly those that are smaller in size, cannot teach everything in the statistics, measurement, and research methodology catalog. Such departments may wisely choose to focus their offerings in research methods in their areas of strength, encouraging their students to partake of complementary classes given by other disciplines on their campus. Consistent with this position, 31% of the responding departments regularly have students take statistics and measurement courses in other departments. Consequently, we inquired as to whether courses taught at an appropriate level were available on campus. As shown in column 5, about half of all respondents indicated that appropriate coverage is available elsewhere on campus. What is not shown is the rate of nonresponse to this question: Approximately 20% of the respondents did not answer the question, often indicating that they did not know what was available outside the department. Nonresponse was just as likely for respondent departments lacking critical courses as it was for those offering the course in question.

Finally, we explored offerings in research methods in more detail (see Table 3). Half of all departments offer a general research methods course, with an average of approximately 85% of the students in these departments taking the course. Approximately 60% of clinical, social, and applied programs; 35% of counseling, developmental, and experimental programs; and 20% of personality programs offer a specialized research methods course. The most pessimistic data in this table are contained in the last column: Over one fourth of departments offer neither a general methods course nor a methods course in the specific area within psychology.

Required Statistics and Measurement Courses

We addressed the duration of basic courses in statistics and measurement required by each of the major program areas in psychology. Respondents were asked to exclude practically oriented clinical assessment courses. With the exception of applied psychology (M = 1.4 years), all of the areas require a mean of 1.1 or 1.2 years. Nearly all (95%) departments offer a universally required doctoral-level introductory statistics course sequence; over 75% of these courses are 1 year in length. Although we did not ask the specific question, we may infer by subtraction

Table 3
Departmental and Disciplinary Offerings in Research Methods

Program area	No. of programs with area	Area offers own research methods course[a]	Department[b]	No methods[c]
Clinical	134	61%	22%	16%
Counseling	22	36	41	23
Developmental	107	39	32	28
Experimental	148	35	28	36
Personality	61	20	31	49
Social	114	57	20	23
Applied	56	61	23	16

Note. In all, 49% of the departments offer a research methods course. Across all departments having such a course, an average of 85% of students take the course.

[a] Includes all programs in a particular substantive area that offer their own research methods courses, whether or not the department also offers a general research methods course. [b] The area does not offer its own research methods course, but the department offers a course that may be taken by all students in the department. [c] Neither the area nor the department offers a research methods course.

that very few of the departments have any requirement in measurement.

Topical Content of Introductory Statistics Sequence

Table 4 summarizes the coverage of several topical areas in the introductory statistics sequence. We asked whether each topical area was covered (a) in depth, so that students could perform the analysis in question themselves; (b) as an introduction, to acquaint students with the concept or the technique; or (c) not at all.

The results indicate that the current course goes little beyond the course as it was taught 20 years ago. The old standards—ANOVA, contrasts and comparisons, and regression analysis—are covered in most departments. But to highlight the lack of progress in incorporating new material, 73% of courses still provide in-depth coverage of repeated measures handled by traditional factorial ANOVA, whereas only 21% provide in-depth coverage of repeated measures handled by multivariate procedures (McCall & Appelbaum, 1973; O'Brien & Kaiser, 1985). Only 20% cover modern exploratory data analysis (Tukey, 1977), and only 18% provide sufficient coverage of power analysis (Cohen, 1977) so that a student could actually calculate the power of his or her own research design. Only 11% of the programs provided in-depth coverage of incomplete designs (e.g., Latin squares), once a major topic of statistics courses in psychology. These results suggest that few of the new developments in statistics are being incorporated into the required statistics sequence and that some of the less central classic topics have been phased out.

Judged Competencies of Graduates to Apply Techniques in Their Own Research

The bottom line of all training is the competency of the graduates. Students may primarily obtain their expertise from informal instruction from their mentors and other graduate students. We therefore asked respondents to judge the competencies of graduates of their program to apply a variety of techniques in their own research. With regard to statistics, Table 5 shows what has been shown over and over. Graduates were judged to be competent at the old standards, but at little that is new. For example, an array of techniques has been developed for the detection and treatment of influential data (e.g., Atkinson, 1985; Belsley, Kuh, & Welsch, 1980; R. D. Cook & Weisberg, 1982), and we have become acutely aware of the extent to which regression analyses can be grossly affected by ill-behaved data. Table 5 shows that over half of the departments said that most or all of their graduates can use ordinary least squares regression, but that 68% of departments judged that few or none of their students know how to detect the influential data points that will bias their regression analyses.

Table 6 paints an even bleaker picture with regard to measurement. Only about one fourth of the departments judged that most or all of their students are competent at methods of reliability and validity assessment; over one third indicated that few or none of their students are. Even smaller percentages of their students are judged to be competent in classical test theory or item analysis. A glance at the remainder of the table indicates that more advanced and newer techniques of measurement and test development are lost on current graduates of psychology programs.

Table 4
Contents of Introductory Statistics Sequence

Topical area	In-depth[a] coverage All[b]	In-depth[a] coverage (Elite)[c]	No coverage All[b]	No coverage (Elite)[c]
Old standards of statistics				
Multifactor ANOVA	73%	(67%)	6%	(0%)
Comparisons	69	(72)	5	(0)
Repeated measures handled by traditional factorial ANOVA	73	(67)	7	(0)
Multiple regression	63	(50)	8	(6)
More advanced considerations				
Exploratory data analysis	20	(22)	33	(22)
Statistical power analysis	18	(17)	27	(28)
Repeated measures handled by multivariate procedures	21	(17)	37	(39)
ANOVA as special case of regression	38	(22)	14	(22)
Analysis of covariance	39	(28)	8	(6)
Multivariate procedures	21	(17)	40	(44)
Incomplete designs	11	(22)	33	(28)
Causal modeling	5	(6)	59	(50)

[a] "In-depth" was defined as coverage to the point that students can perform the analysis in question themselves. [b] Percentage of all programs with introductory statistics sequences. [c] Percentage for the 21 elite programs responding to the survey.

Table 5
Judged Competencies of Graduates to Apply Techniques of Statistics in Their Own Research

Technique	Percentage of programs indicating whether graduates can apply techniques to their own research: Most or all[a] (≥75%) All[c]	Most or all[a] (≥75%) (Elite)[d]	Few or none[b] (≤25%) All[c]	Few or none[b] (≤25%) (Elite)[d]
Old standards of statistics				
Multifactor analysis of variance (ANOVA)	81%	(71%)	4%	(5%)
Contrasts/comparisons	76	(71)	6	(5)
Repeated measures by factorial ANOVA	74	(67)	5	(5)
Ordinary least squares regression	58	(48)	6	(0)
Basic nonparametric procedures	48	(33)	20	(10)
Statistical procedures in common use				
Repeated measures by multivariate procedures	22	(19)	37	(19)
Analysis of covariance and alternatives	38	(33)	22	(29)
Multivariate analysis of variance	18	(10)	34	(24)
Other multivariate procedures	11	(0)	51	(44)
Alternatives to ordinary least squares regression	3	(5)	78	(52)
Data treatment procedures				
Modern graphical data display	15	(14)	55	(57)
Data transformations	31	(29)	30	(24)
Detection and treatment of influential data	8	(5)	68	(71)
Causal modeling and other exotics				
Path analysis	2	(0)	81	(86)
Confirmatory factor analysis	3	(0)	72	(67)
Latent variable structural models	2	(0)	81	(81)
Time series analysis	1	(0)	86	(90)

[a] Percentage of programs that judged that most or all of their students could use the technique in their own research. [b] Percentage of programs that judged that few or none of their students could use the technique in their own research. [c] All programs responding to survey. [d] 21 elite programs responding to the survey are in parentheses.

Finally, what of design? The picture, given in Table 7, provides no respite. Most graduates were judged able to design laboratory experiments. But what about the person by situation design, the frequently used design in which a subject classification variable is crossed with experimental variables? One third of all departments indicated that few or none of their graduates could implement this design in their own research. Psychologists are increasingly asking questions that require the study of phenomena over time. Yet competence at over-time designs is apparently very low. Applied research is increasingly being incorporated into the core of psychology and is a rich source of jobs for graduates; yet, these same graduates are not judged to be competent at quasi-experimentation. An increasing number of review articles use meta-analytic techniques; yet, few of our graduates are judged to be competent in the use of these techniques.

Comparison of Course Areas Studied by Practitioner-Oriented Versus Nonpractitioner-Oriented Students

The findings presented thus far are discouraging. However, one alternative possibility is that training is as good as ever in the core subdisciplines of psychology, but that the growing number of clinical and practitioner students receive substantially less training. We contrasted the training of practitioner- versus nonpractitioner-oriented students in formal research methods course work. Table 8 summarizes our findings. Course areas refer to those in which students receive at least half a semester of coverage. Columns 1 and 2 of Table 8 address clinical and practitioner-oriented students; columns 3 and 4 address nonpractitioner-oriented students. A quick review of the table indicates that the two groups of students have very similar curricula.

Comparison of Elite Versus Other Graduate Programs

A second possible reason for the present discouraging findings is that a decreasing percentage of PhD recipients

Table 6
Judged Competencies of Graduates to Apply Techniques of Measurement in Their Own Research

Technique	Percentage of programs indicating whether graduates can apply techniques to their own research: Most or all[a] (≥75%) All[c]	Most or all[a] (≥75%) (Elite)[d]	Few or none[b] (≤25%) All[c]	Few or none[b] (≤25%) (Elite)[d]
Old standards of measurement				
Classical test theory	19%	(10%)	53%	(67%)
Exploratory factor analysis	12	(0)	45	(33)
Item analysis	17	(5)	57	(67)
Methods of reliability measurement	27	(14)	38	(52)
Methods of validity measurement	22	(10)	44	(52)
Scaling procedures				
Unidimensional scaling	5	(10)	69	(62)
Multidimensional scaling	2	(0)	74	(52)
More advanced developments				
Item response theory	6	(0)	76	(81)
Generalizability theory	6	(5)	75	(71)
Selection models	3	(0)	80	(86)
Bias analysis	1	(0)	89	(76)
Equating	2	(0)	87	(81)

[a] Percentage of programs that judged that most or all of their students could use the technique in their own research. [b] Percentage of programs that judged that few or none of their students could use the technique in their own research. [c] All programs responding to survey. [d] 21 elite programs responding to the survey.

Table 7
Judged Competencies of Graduates to Use a Variety of Designs in Their Own Research

Technique	Percentage of schools indicating whether graduates can apply designs to their own research: Most or all[a] (≥75%) All[c]	Most or all[a] (≥75%) (Elite)[d]	Few or none[b] (≤25%) All[c]	Few or none[b] (≤25%) (Elite)[d]
Basic designs				
Laboratory experiments	83%	(90%)	35%	(0%)
Field experiments (basic research in field settings)	42	(48)	13	(14)
Experimental personality (person by situation design)	25	(33)	34	(29)
Over-time designs				
Longitudinal designs	12	(5)	52	(57)
Concommitant (multivariate) time series	2	(0)	90	(90)
Single-subject designs	10	(0)	59	(76)
Experimental and quasi-experimental designs for applied problems				
Large-scale social experiments	7	(5)	55	(62)
Regression discontinuity	4	(0)	82	(81)
Interrupted time series	5	(0)	79	(86)
Nonequivalent control group	12	(0)	58	(71)
Structural equation methods				
Path-analytic designs	3	(5)	75	(71)
Latent structure (multiple indicator)	2	(0)	84	(86)
Qualitative methods	5	(5)	68	(62)
Meta-analysis	5	(5)	80	(91)

[a] Percentage of programs that judged that most or all of their students could use the design in their own research. [b] Percentage of programs that judged that few or none of their students could use the design in their own research. [c] All programs responding to survey. [d] 21 elite programs responding to survey.

in psychology are being trained in "elite" institutions (A. Howard et al., 1986). Our elite institutions may continue to foster rigorous training, whereas our less established training programs may not. We compared the available responses (21) received from the elite departments with those from the lower ranked departments. The response rate for the elite departments (87.5%) did not differ from the overall response rate (84%) for all departments.

Table 1 paints the expected picture of greater available faculty resources in the elite programs for training in statistics and measurement than is found in the sample of all programs. The elite programs have larger faculties, more first-year graduate students, and more faculty whose training and instructional focus are in the quantitative area. Interestingly, a larger percentage of the elite programs (48%) relative to the percentage in the total sample (31%) regularly have PhD students take statistics or measurement courses in other departments.

In contrast, a comparison of the requirements and the outcomes of training suggests a more mixed picture. About the same percentage of the elite programs as in the total sample offer an introductory graduate statistics sequence (86% vs. 89%) that is typically required of all PhD students (89% vs. 93%) and that is typically one year in length (78% vs. 77%). Table 2 shows that a greater percentage of the elite schools relative to the total sample offer at least a one-half semester course in scaling, mathematical psychology, and causal (structural equation) modeling. In contrast, a smaller percentage of the elite schools offer at least a one-half semester course in evaluation research and test construction. The elite schools do not differ from the total sample in their topical coverage of the introductory statistics sequence (see Table 4). Finally, as can be seen in Tables 5, 6, and 7, there is little difference between the percentage of elite versus the total sample of programs in the judged competencies of their graduates to apply techniques of statistics, measurement, and design to their own research. Thus, it does not appear that there are substantial differences in the training of the average student in statistics, measurement, and methodology at the elite versus the other schools. It is likely that the benefits of the greater resources of the top schools are limited to a relatively small number of students who specifically focus their graduate studies on quantitative methods and methodology.

Table 8
Course Areas Taken by Practitioner-Versus Nonpractitioner-Oriented Students

Course area[a]	Clinical or other practitioner-oriented students (%): Most or all (≥75%)	Clinical or other practitioner-oriented students (%): Few or none (≤25%)	All nonpractitioner-oriented students (%): Most or all (≥75%)	All nonpractitioner-oriented students (%): Few or none (≤25%)
Old standards of statistics				
Analysis of variance	95%	3%	93%	5%
Multiple regression	72	14	70	13
Multivariate analysis	36	35	38	25
Factor analysis	18	59	20	51
Measurement				
Test construction	20	64	9	76
Scaling	11	84	11	72
Methodology				
Evaluation research				
Quasi-experimental	9	72	6	75
design	24	58	17	60
Survey research	11	80	6	79
Other topics				
Computer applications				
Mathematical	44	37	58	21
psychology	1	98	2	88
Causal modeling	7	82	7	78
Time series	6	88	6	86

[a] Course areas in which students receive at least one-half semester of coverage.

Commentary

Before we consider some of the implications of and remedies for the failure to advance in the quality of training in statistics and methodology, coupled with the decline in the quality of training in measurement, it is useful to consider the perspectives of several core areas of psychology on training in research methods. Areas may differ greatly in their training needs. Considering briefly the example of experimental psychology, it is possible that this area may have little need for anything other than classic training in laboratory experiments and analysis of variance. Few of the analytic and methodological techniques included in the survey, such as structural equation models, are currently represented in experimental journals (but see, e.g., Geiselman, Woodward, & Beatty, 1982; Sekuler, Wilson, & Owsley, 1984). In contrast, other techniques not fully represented in the present survey (e.g., mathematical models) may be critical to this area. Other substantive areas may also have a greater or lesser perceived need for the new developments in research methods. To provide a better understanding of the perspectives of several diverse areas of psychology, we solicited commentary from four senior investigators who are current or recent editors of APA journals. Each of these individuals was presented with a preliminary draft of the results of this survey. They were asked to briefly discuss how their own substantive area is changing in terms of its research questions and approaches. They were also asked to address the needs of their area in terms of training and practice in research methods. The commentary of each of these individuals is presented below.

Experimental Psychology: Henry L. Roediger III

The survey provides important information on how well graduate departments in psychology are educating their students in quantitative topics. The overall news is not particularly good, with little in the way of formal course work provided. A good case can be made that education in quantitative topics has not kept pace with the increased sophistication of the field. My commentary will concentrate on the implications for experimental psychology, particularly human experimental psychology.

1. Although the formal course work on quantitative topics has not increased, cries of alarm may be premature. Much graduate education occurs under the tutelage of a faculty member, and thus much quantitative education in particular fields probably occurs by this more informal means.

2. Even so, the standard graduate sequence in statistics cannot be counted on to educate students adequately on quantitative topics. Such courses must be supplemented. For cognitive psychologists, the most obvious candidates are courses in mathematical psychology, mathematical models, and computer simulations. Such courses are probably already offered in leading cognitive psychology programs.

3. Another possible supplement for students being trained in human experimental psychology would be a required course in psychophysics. As far as I can tell, training in this venerable branch of experimental psychology is virtually nonexistent. Yet, if required to take a good psychophysics course, students would be exposed to many of the thorniest problems in measurement. In addition, they would come to appreciate the lost art of small-N design in human experimental psychology.

4. An obvious candidate to improve the quantitative skills of our students is to require more mathematics courses as a prerequisite to graduate school. Alternatively, if bright students are admitted without such courses, these might be required upon their admission. There seems little doubt that improvements in quantitative education in North American psychology departments will be required to keep abreast of the field.

Developmental Psychology: Sandra Scarr

When most developmental research consisted of experimental studies of small samples in laboratory settings, the analysis of variance and the t test sufficed, in most professors' minds, to give students the statistical tools of the trade. Today most developmental research concerns naturally occurring behaviors in natural settings in which the experimenter's control over "independent" variables and random assignment to conditions is essentially nil. Unfortunately, most instruction in quantitative methods continues to emphasize limited views of sampling theory, to base analyses on improbable distributions, to encourage ignorance of effect size in favor of arbitrary p values, and largely to ignore the intrinsically confounded nature of many important developmental phenomena.

Quantitative methods that support contemporary developmental research must stress regression and correlation-based analyses, whether in structural equation modeling, time series analysis, or multidimensional scaling. Although important experimental research continues in areas where traditional analyses can apply, far more developmental research today requires nonexperimental modeling of complex, existing phenomena that are not under the investigator's control. Unfortunately, students and some of their mentors seem to feel quite comfortable in analyzing data by assuming that people are randomly assigned to their niches and that a measured attribute is independent of all other possible unmeasured attributes. Appropriate instruction in developmental quantitative methods must also take into account change over time, the very essence of developmental concerns. Despite repeated attempts by McCall, Appelbaum, Wilson, Kenny, and others to illuminate the analyses of change, it seems that the issues are seldom being taught to developmental students.

Perhaps it is unrealistic to hope that most developmental students will gain sophistication in the most advanced quantitative models, given the demands of their substantive curriculum. One could at least hope for an inquisitive posture that would lead them to seek appropriate consultation. Such curiosity will depend on instruction that instills an attitude about the importance of effect sizes, the variety of underlying probability distributions, the correlated nature of most real-world phenomena, and the nonrandomness of life.

Social Psychology: Steven J. Sherman

The survey assesses the extent to which the many recent developments in measurement, statistics, and methodology have been incorporated into PhD programs in psychology. Not surprisingly, the requirements and course offerings have lagged far behind the rapidly increasing quantitative sophistication in the field. This is probably due in part to two factors. First is the problem of finding current faculty or hiring new faculty to teach students these recent developments. Second is the tendency for programs to stress more and more research productivity. Academic job possibilities for a recent PhD recipient are bleak without a solid publication record. Many programs have thus chosen an "immersion in research" approach from the very first semester of graduate training. Ironically, this stress on research and publications comes at the expense of content courses and training in quantitative skills and methodology. Thus, students are encouraged to achieve a high level of research productivity without much consideration for training in the content, methodological, and quantitative skills necessary for good research.

On the other hand, programs do recognize that formal training and course work are not the only ways in which methodology and statistics can be learned. Thus, there is an increased tendency toward informal ways of learning the quantitative and methodological skills that are necessary to any specific research endeavor. Given motivated students and given that the necessary knowledge is available from sources in the program, this is not an unreasonable approach. After all, many faculty who were not trained in current measurement, statistical, and methodological approaches have had to learn them nevertheless.

The fact remains that current graduate education in quantitative topics is not sufficient for turning out solid and competent researchers and scientists. With regard to social psychology in particular, there have been many new directions in research that require certain kinds of statistical and methodological training that are not now available or not encouraged in many graduate programs. First, there is a good deal of recent research (especially in social cognition) that makes use of models and computer simulations. Second, many areas of social psychology have begun to use a decision-making point of view. Normative models, Bayesian statistics, and advanced regression approaches are very important in such work. Third, applied social psychological work has turned more and more to longitudinal studies. Structural equation approaches are absolutely necessary for such work; yet, as the survey points out, competence at over-time designs is very low. Fourth, areas of social psychology that try to integrate the roles of personality and situational factors in behavior require expertise in the person by situation design. Finally, given the growing number of studies in various areas of social psychology, it has become important to look for effects across studies. Meta-analytic approaches have been developed to resolve such questions.

In the case of each of these new directions, appropriate training in the relevant statistical and methodological procedures is not available in our graduate programs. It is neither feasible nor desirable that all our students be trained in all these approaches. Also, it is not necessarily a disaster that our programs have lagged behind current and recent advances in measurement, statistics, and methodology. Yet, it *is* important that such training be available for our students (and for our existing faculty), and it *would* be a disaster if our graduate programs did not make a concerted effort to hire faculty or to otherwise ensure that such advances be represented.

Clinical Psychology: Alan E. Kazdin

Research in clinical psychology draws on an extraordinary range of research designs. Apart from the relatively straightforward randomized experimental investigation, longitudinal-prospective, cross-sectional, and case control studies and a wide array of quasi-experimental designs are used routinely to address questions about clinical dysfunctions. The very scope of designs that are used and the complexity of naturalistic (rather than experimentally manipulated) variables that are studied require special knowledge of methodology. Once an investigation is under way, a number of other characteristics are likely to emerge that can interfere with drawing valid inferences. Thus, in clinical settings, missing data for a few of the multiple measures obtained on individual subjects and complete loss of data for some subjects (i.e., attrition) are relatively common. These and related problems require special training because they need to be addressed at the design and evaluation stages.

In addition to design, statistical evaluation raises issues that further argue for the need for extensive training in methodology. First, inappropriate data analyses are still relatively common. Confusion over when and how to use multivariate analyses and factor analyses, for example, often leads to inappropriate analyses in clinical research. Second, the full range of valuable statistical information for a particular data set is overlooked. Studies focus almost exclusively on tests of significance to compare means among groups. Examination of a broader range of questions from a set of data (e.g., effect size and confidence limits) is relatively rare. Consequently, studies are quite restricted in the picture they provide of the data. Finally, several analyses (e.g., path analysis, confirmatory factor analysis, and survival analysis) are available to address many of the questions of interest in clinical research. These analyses remain relatively esoteric from the standpoint of clinical research even though they are suited to test models of clinical dysfunction, develop assessment tools, and examine the long-term impact of treatment. A range of analyses out of the mainstream of most graduate training programs remains to be incorporated into areas of research that would profit remarkably from their use.

Intensive training in methodology, design, and statistics is especially important in clinical psychology given the range of questions that are addressed and the situations in which clinical researchers function. Perhaps even more than addressing the many design questions that such training affords, there is a type of thinking that one hopes to instill from training in methodology. Methodology and statistical analyses, as well, provide the researcher with models of how to conceptualize, explore, and evaluate phenomena. The thinking underlying training in methodology, design, and statistics may be just as important as the content that is conveyed.

Conclusions: Implications, and Remedies

The results of the survey show that new developments in statistics, measurement, and methodology are not being incorporated into most graduate training programs. A similar finding in any *substantive* area in psychology that new developments were not being incorporated into graduate training programs would have grave implications for the field of psychology. However, the implications of the present findings in the broad area of research methods are more multifaceted, because statistics, measurement, and methodology are regarded by most psychologists simply as tools for conducting research in their own substantive area rather than as important substantive areas in their own right.

Conclusions From the Survey and Commentaries

Five conclusions stand out from a consideration of the results of the survey and the commentaries.

1. *Current PhD students are receiving traditional training in methodology and statistics, training that primarily supports laboratory rather than field research.* The commentaries by Roediger, Scarr, Sherman, and Kazdin indicate that the implications of the survey's results vary by substantive fields within psychology. Academic experimental psychologists and laboratory social psychologists continue to emphasize problems that can be well addressed by factorial designs conducted in the laboratory and analyzed with analysis of variance,[3] topics that are well addressed in the standard curriculum. In contrast, developmental, applied social, and clinical psychologists have turned their attention to a variety of new substantive questions that address issues in naturalistic settings outside the laboratory, with specific populations of interest, studied over time, and often without the benefits of random assignment to treatment conditions. Thus, at the end of the first-year graduate statistics sequence, students working in laboratory settings have a strong foundation for handling many of the design and analysis issues they will later confront in their substantive research. Students working in field settings have received the same standard training, but the techniques are rarely applicable to their research settings, leaving them essentially unprepared to face the design and analysis issues that will confront them in their substantive research.[4]

2. *Even with "ideal training" in a first-year graduate sequence, supplementary training is required.* Even if an ideal first-year training sequence were implemented (see Remedy 1 below), supplements to the curriculum would be required. For research in laboratory experimental psychology, training in such areas as psychophysical approaches, in mathematical modeling, in decision theory, and in computer simulation would be valuable for certain research endeavors. For research in field settings, a large number of topics, such as quasi-experimental designs, longitudinal designs, advanced regression techniques, and structural equation models, should be considered, depending on the needs of the specific substantive area.

3. *Measurement has declined substantially in the curriculum.* Measurement-related issues continue to permeate the discipline from psychopathology through psychophysics. Yet, students are often unacquainted with even the classic concepts that underlie basic psychological measurement. This deficiency opens the door to a proliferation of poorly constructed ad hoc measures, potentially impeding future progress in all areas of the field. Perhaps less troubling, at least at this time, is the near total lack of attention to more specialized areas of modern measurement theory, such as item response theory, whose usefulness for constructs other than abilities is just beginning to be explored (Thissen & Steinberg, 1988).

4. *Training in new techniques and methodologies is generally unavailable within the psychology curriculum.* The results of the survey strongly indicate that new advances have not been integrated into or added to the curriculum. The situation is analogous to limiting formal training of students in substantive areas of the discipline to material appearing in journal articles published through the mid-1960s. We revise substantive curricula yearly and frown upon colleagues who do not keep their courses up to date. Yet, the quantitative/methodological curriculum, in the main, has remained virtually static over the years.

5. *There is a substantial lack of awareness about other resources on campus that may provide training for students, even though such training is sorely needed.* Although not true at all universities, many programs fail to use formal training opportunities available in other departments on campus. We have found useful quantitative curricula in economics (econometrics), sociology (survey research, survey sampling, and causal modeling), mathematics (time series analysis and advanced regression), business (exploratory data analysis, panel designs, and categorical data analysis), and education (measure-

[3] Many experimental psychology graduates currently obtain positions in business and industry (see A. Howard et al., 1986). In many such settings, traditional training provides poor preparation for the types of methodological and analysis issues that arise.

[4] Ironically, many programs, even those in developmental, applied social, and clinical psychology, encourage graduate students to conduct laboratory experiments for their thesis and dissertation research projects. Students, in turn, may feel too ill prepared in modern methodology, measurement, and statistical analyses to resist this encouragement and to take on and defend a complex field research project before a committee of critical, traditionally trained faculty members. Thus, traditional training appears to provide students with excellent training for traditional theses and dissertations, but not for the research that many of them will conduct upon completion of their PhDs.

ment theory). One way in which the elite schools stand out from others is in their use of course work in other departments.

Implications of Deficient Training for the Discipline

In interpreting the results of the survey, it is particularly important to give careful consideration to the costs of *not knowing* a particular statistical, measurement, or methodological concept or approach. Broadly stated, these costs are three in number: (a) failure to use those designs that optimize the potential tests of theory, (b) failure to gather those data that would best inform the research questions at hand, and (c) failure to reach correct conclusions because of misanalysis of data.

These costs are not equal in detriment to the discipline. It is less important, for example, for researchers to know how to conduct time series analyses themselves than it is for them to know that the data collected over time or geographical location often introduce problems of nonindependence that require specialized analytical techniques. It is less important that investigators be able to conduct survival (event history) analyses themselves than to understand that time to failure (e.g., the length of time [if ever] following treatment before heroin addicts or depressed clients revert to their previous status) is a potentially powerful outcome variable that should be collected in certain designs and that is analyzed using specialized techniques. Finally, as noted by Scarr and Kazdin, some developments are critical because they dramatically alter the ways in which we conceptualize phenomena and conduct research. For example, randomized factorial experiments and the use of ANOVA foster a conception of variables as representing independent and orthogonal dimensions, a conception that dominates theorizing in many areas of psychology. The development of quasi-experimentation has fostered new ways of thinking about design and analysis, notably critical multiplism (T. D. Cook, 1985; Houts, Cook, & Shadish, 1986), in which multiple comparison groups, multiple designs, and multiple analyses having different biases are developed to determine the degree to which they triangulate on the same result. Structural equation modeling has led to a new emphasis on the links between substantive theory, measurement, design, and analysis and the development of hierarchical sets of models that are competitively tested. This final group of techniques clearly requires a more in-depth presentation to ensure that graduate students are exposed to the important conceptual ideas underlying the full range of methodological and quantitative approaches applicable to their substantive area. Nonoptimal conceptualization leads to improper designs or improper or incomplete data collection that cannot be remedied. Nonoptimal analyses, in contrast, can be recomputed.

Deficiencies in quantitative and methodological training do have negative implications for the progress of substantive areas of the discipline. These deficiencies obviously can lead to seriously flawed research efforts. Were these flawed research efforts inevitably detected at the editorial review stage prior to publication, psychology would pay little penalty except in terms of the wastefulness of many of the efforts of its researchers. However, a review of traditional designs and analyses in several of the leading APA journals (Sechrest, 1987) indicated that shortcomings in data quality abound, such as inattention to reliability and validity of measurements, inappropriate statistical analyses, inattention to issues of statistical power, and unjustified interpretation of findings. Such findings suggest that the weakness of quantitative and methodological training in the discipline means that reviewers themselves may frequently miss grave difficulties in manuscripts submitted for publication. These problems can be magnified with newer, less well-known techniques. For example, complex structural equation models can be underidentified, meaning that the structural coefficients from which causal inferences are made are in reality random numbers. Editors increasingly find themselves with substantive reviews strongly supporting publication coupled with a methodological review pointing out that the manuscript is fatally flawed. If the discipline does not become better educated at the quantitative/methodological aspects of current research, serious errors may become increasingly prominant in our published literature.

Remedies

Before proposing remedies, we need to recognize that there is simply not sufficient room in a 4-year curriculum to teach students all or even most of the new developments in statistics, methodology, and measurement. Nor are current university budgets so bountiful that every PhD program in psychology can expect to hire a cadre of quantitative specialists who would create a complete curriculum of new developments. Within these limitations, however, we see five broad strategies that may help remedy the current situation.

1. *The quantitative and methodological curriculum should be revamped carefully to be sensitive to current needs.* We should no longer assume that each graduate student must be trained to perform all of his or her own analyses. Rather, we should assume that what students require is *proficiency* in those techniques that they are most likely to use and *acquaintance* with the basic conceptual ideas of a variety of techniques that may be useful in research questions they may later encounter. Content analysis of the research methods and analyses reported in current journals can provide information about current practice. Joint discussions of statisticians, methodologists, and researchers in a substantive area can help define some of the research questions that will drive future methodological and analytic needs. Such information would be very useful in designing model introductory and more advanced courses that would adequately represent the new developments of particular importance to each of the substantive areas.

Although we believe such a full review of the statistics curriculum is desirable, two readers of a preliminary version of this article suggested ideas for revamping the basic graduate statistics sequence. Peter Bentler (personal communication, January 19, 1989) suggested that the required undergraduate preparation in mathematics and statistics be strengthened.[5] Then the lowest level review material could be dropped from the graduate sequence, and the traditional statistics sequence of analysis of variance and multiple regression could be covered in one semester or two quarters. Thereafter, students would opt into a two-semester or two-quarter sequence covering either experimentally relevant methods or those methods and measurement issues more relevant to field research. In contrast, Sanford Braver (personal communication, February 10, 1989) argued for a sequence in which all students take multiple regression and most students take an analysis of variance course because of its frequency of use in journals in most areas of psychology. All students would then take a survey course in which they develop an acquaintance but not proficiency with as many of the other classic and newer areas of statistical analysis as possible. It would be particularly important in this course to provide resources describing available instruction on campus and important references for self-instruction in each of the techniques. Although these two suggestions by no means exhaust the possible ways in which the statistics sequence can be revised, they do serve as important illustrations that can help stimulate critical analysis of this issue. Other particularly important issues are the role of measurement in the curriculum of both experimental and nonexperimental students, the extent to which a statistics curriculum will need to be individually tailored depending on each student's substantive problem area, and the way in which training in statistics, measurement, and methodology should be integrated within the course sequence.

2. *Available training opportunities across campus should be culled carefully.* This suggestion is not tantamount to scanning a graduate catalog for courses with appropriate titles. Psychology faculty need to pursue potential courses actively, meeting with their colleagues who teach the courses to discuss course goals versus student needs, interviewing graduate students who have taken the course, and even actually sitting in on the course to assess its quality and level of instruction, as well as the background assumed for the course. In addition, courses in other departments will have different orientations and will draw examples from different disciplines (e.g., in econometrics, nonrecursive [reciprocal] relationships will be taught with examples from supply and demand; in sociology, with examples of economics and labor force participation). The concepts taught in these courses need to be brought home to psychology, perhaps through "laboratory" practice sessions led by faculty of psychology, in which students are exposed to examples of the new methodologies that use data more central to psychology itself.

3. *Students should be trained in informal settings.* Both Roediger and Sherman noted the importance of informal training in quantitative analyses and methodology. Nonetheless, programs and even professors within a program differ in the amount of informal training in research methods that they offer. Innovative new approaches that permit and encourage students to participate in raising and solving problems in research methods are needed. Many programs currently hold "brown bag" or other informal meetings, usually focused on the design of laboratory research projects. Broadening these meetings so that they consider a wider array of methodological approaches would be a potentially useful step. Supplementing these meetings with data-analytic brown bags (Rosenthal, 1987) in which difficult data analysis problems are discussed by researchers and one or more experts in statistics, whether from psychology or from other departments, would help ensure students' exposure to many of the new developments in statistics and measurement. In our experience, to be most useful, participants must be provided with appropriate background reading to introduce the session topic, and the sessions themselves must be conducted in a nonthreatening, constructive style. Under these conditions, such informal sessions serve not only as a valuable training experience for graduate students and faculty alike, but they also help improve the quality of publications coming out of the department.

4. *Faculty who may have fallen behind should be retrained.* Successful implementation of the first three remedies would require that a significant portion of the faculty be knowledgeable about the new methodological and quantitative approaches relevant to their areas to provide the necessary formal and informal training. Psychology would require the development of postdoctoral, intensive summer, and within-university faculty development programs in quantitative techniques and methodology for those individuals who wish to expand their areas of expertise or keep up to date with new developments in statistics, measurement, and methodology related to their substantive areas of research interest. A long existing example of such training is the University of Michigan Summer Quantitative Study program. Interestingly, this social science–oriented program has been dominated on both the faculty and student side by other social sciences (e.g., sociology and political science) rather than psychology. APA does typically offer one or two continuing education courses on quantitative topics at its annual conventions. This situation contrasts sharply with the extensive offerings available in clinical psychology.

5. *Methodological and statistical review of substantive manuscripts that use advanced techniques should be sought.* Substantive journals not only archive findings, but also serve as repositories of the acceptable methodological and statistical practices in the area. As such, journal editors need to exercise special care in the review of articles using advanced techniques. Methodological/statistical reviews need to be sought in addition to substantive reviews to ensure that techniques have been properly used and results properly interpreted. Such reviews reduce the likelihood that inappropriate applications of "state of the art" methodologies will be published,

[5] Difficulties will occur in implementing this proposal because of the declining pool of college students in the United States who have a strong background in mathematics and statistics. As one index of this decline, the percentage of college freshmen planning to major in mathematics dropped from 4.6% in 1966 to 0.6% in 1988 (Green, 1989).

becoming in time blindly followed models for the use of the technique in the area.

Concerns About Our Human Capital

In order to remedy deficiencies in quantitative and methodological training, psychology will need to develop and nurture its talent pool. If students and faculty are to learn to recognize situations in which complex methodological and statistical issues arise; if informal, brown bag meetings that discuss general issues in research methods and data analysis are to be offered; or if faculty development in statistics, measurement, and methodology is to occur, then there will need to be individuals who have the expertise to lead these developments.

Our first concern is that within the current structuring of psychology departments, there may not be the necessary resources for broad program enhancement. Only 17% of all departments responding to the survey offer a PhD in a quantitative area; a third of the quantitative programs had no new first-year students. The quantitative areas range widely in the concentrations offered from mathematical psychology (n = 6), to applied statistics (n = 15), to psychometrics and measurement (n = 9), to quantitative programs associated with a particular substantive area such as personality (n = 5). At the time of this survey, a total of 108 students were being trained in these quantitative programs. With the usual loss rates from graduate school and the strong competition for quantitative graduates from industry and research organizations, this poses a bleak picture of the number of new PhDs with strong quantitative training who will be available to replenish our greying academic quantitative force. The average age of the quantitative force is high; in 1985 the median age of all members of Division 5 of APA (Evaluation, Measurement, and Statistics) was 51; for fellows, it was 65.

A further human capital limitation arises from the politics of psychology departments with respect to their quantitative faculty. Following a distinction made by Muthén (1989), we can identify two classes of quantitative faculty. First are the high-level "developers," the small number of individuals who publish highly technical articles that develop original methodologies and statistical and measurement techniques. When not housed in a quantitative area, these individuals are at considerable risk for being denied promotion and tenure because their work is difficult to evaluate, it is not viewed as central to psychology (cf. Campbell, 1969), and they are often uninterested in addressing the everyday consulting concerns of their substantive colleagues.

A second group are the "bridgers," who are well trained quantitatively but do not publish original methodological or quantitative work. These bridgers are typically individuals trained in a substantive area who have pursued quantitative training beyond that required for their substantive degrees. They are perceived to be excellent hires for the department, who feel that they are getting a real bargain by hiring someone who can contribute in a major way to a substantive area while serving the quantitative training and consulting needs of the department. Unfortunately, like many apparent bargains, this one often proves under the current system of faculty evaluation to be costly both to the department and to the individual who is hired. If the member meets the consulting needs of the department (which are likely to be enormous by the time a department perceives the need to hire a quantitative faculty member), then the faculty member is unlikely to have sufficient time to simultaneously make major strides in the substantive area. And it is on the substantive contributions that the faculty member will almost certainly be evaluated for promotion and tenure. Joint publications of bridgers with other substantive faculty are typically viewed as part of the substantive faculty member's research program, with the bridger being merely the data analyst. Hence, these important bridgers are often doomed to failure in their quest to serve two masters. Under the press of becoming tenured, they may be forced to withdraw from the consulting role and to stop keeping up with new quantitative/methodological advances, thereby thwarting the very reason, from the department's perspective, for which they were hired. Departments of psychology will have to carefully consider their values and expectations when they hire the bridgers they so badly need. Our proposed model of quantitative/methodological attainment for PhDs, which assumes competence with the most common approaches in each substantive area, but only acquaintance with other approaches that may be needed in the research, strongly presumes the availability of bridgers who can perform the up-to-date, high-level consulting that is required.

In sum, psychology needs to focus its attention on the quantitative and methodological aspects of its training. A concerted effort will be needed to strengthen programs. Absent this effort, psychology as a discipline will be limited in its progress by its outdated tools.

REFERENCES

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit and probit models.* Beverly Hills, CA: Sage.

Atkinson, A. C. (1985). *Plots, transformations and regression.* New York: Oxford University Press.

Belsley, D. A., Kuh, E., & Welsch, R. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity.* New York: Wiley.

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

Bishop, Y. M., Feinberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (rev. ed.). San Francisco: Holden-Day.

Campbell, D. T. (1969). Ethnocentrism of disciplines: The fish-scale model of omniscience. In M. Sherif & C. W. Sherif (Eds.), *Interdisciplinary relationships in the social sciences* (pp. 328–348). Hawthorne, NY: Aldine.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression.* New York: Chapman and Hall.

Cook, T. D. (1985). Postpositivist critical multiplism. In R. L. Shotland & M. M. Mark (Eds.), *Social science and social policy* (pp. 21–62). Beverly Hills, CA: Sage.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.

Coombs, C. (1964). *A theory of data.* New York: Wiley.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wiley.

Duncan, O. D. (1975). *Introduction to structural equation models.* New York: Academic Press.

Geiselman, R. E., Woodward, J. A., & Beatty, J. (1982). Individual differences in verbal memory performance: A test of alternative information processing models. *Journal of Experimental Psychology: General, 111,* 109–134.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Green, K. C. (1989). A profile of undergraduates in the sciences. *American Scientist, 77,* 475–480.

Greenacre, M. J. (1984). *Theory and applications of correspondence analysis.* New York: Academic Press.

Gulliksen, H. (1950). *Theory of mental tests.* New York: Wiley.

Hambleton, R. K., & Swaminathan, H. (1984). *Item response theory: Principles and applications.* Boston: Kluwer-Nijhoff Publishers.

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

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Houts, A. C., Cook, T. D., & Shadish, W. R., Jr. (1986). The person-situation debate: A critical multiplist perspective. *Journal of Personality, 54,* 52–105.

Howard, A., Pion, G. M., Gottfredson, G. D., Flattau, P. E., Oskamp, S., Pfafflin, S. M., Bray, D. W., & Burstein, A. G. (1986). The changing face of American psychology: A report from the committee on employment and human resources. *American Psychologist, 41,* 1311–1327.

Howard, G. S., Cole, D. A., & Maxwell, S. E. (1987). Research productivity in psychology based on publication in the journals of the American Psychological Association. *American Psychologist, 42,* 975–986.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jones, L. V., Lindzey, G., & Coggeshall, P. E. (1982). *An assessment of research-doctorate programs in the United States: Social sciences.* Washington, DC: National Academy Press.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge, MA: Abt Associates.

Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions.* New York: Cambridge University Press.

Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis: Models of quantitative change.* New York: Academic Press.

Lord, F. M. (1980). *Applications of item response theory to practical test problems.* Hillsdale, NJ: Erlbaum.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated-measures designs: Some alternative approaches. *Child Development, 44,* 401–415.

McCleary, R., & Hay, R. A. (1980). *Applied time series analysis for the social sciences.* Beverly Hills, CA: Sage.

Muthén, B. (1989). The future of methodological training in educational psychology: The problem of teaching students to use new sophisticated techniques. In M. Wittrock & F. Farley (Eds.), *The future of educational psychology* (pp. 181–189). Hillsdale, NJ: Erlbaum.

Nesselroade, J. R., & Baltes, P. R. (Eds.). (1979). *Longitudinal research in the study of behavior and development.* New York: Academic Press.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97,* 316–333.

Riecken, H. W., Boruch, R. F., Campbell, D. T., Coplan, W., Glennan, T. K., Pratt, J., Rees, A., & Williams, W. (1974). *Social experimentation: A method for planning and evaluating social innovations.* New York: Academic Press.

Rosenthal, R. (1987, August). Methodological spirit and methodological substance in social psychology. In L. S. Aiken & S. G. West (Chairs), *Adequacy of methodological and quantitative training: Perspectives of the disciplines.* Symposium conducted at the meeting of the American Psychological Association, New York.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling: Theory, methods and applications.* New York: Academic Press.

Sechrest, L. (1987, August). Data quality: The state of our journals. In L. S. Aiken & S. G. West (Chairs), *Adequacy of methodological and quantitative traing: Perspectives of the disciplines.* Symposium conducted at the meeting of the American Psychological Association, New York.

Sekuler, R., Wilson, H. R., & Owsley, C. (1984). Structural modeling of spatial vision. *Vision Research, 24,* 689–700.

Shepard, R. N., Romney, A. K., & Nerlove, S. B. (Eds.). (1972). *Multidimensional scaling: Theory and applications in the behavioral sciences.* New York: Seminar Press.

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1–49). New York: Wiley.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science, 133,* 80–86.

Thissen, D., & Steinberg, L. (1988). Data analysis using item response theory. *Psychological Bulletin, 104,* 385–395.

Torgerson, W. S. (1958). *Theory and methods of scaling.* New York: Wiley.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Winer, B. J. (1962). *Statistical principles for experimental design.* New York: McGraw-Hill.