Doctoral Training in Statistics, Measurement, and Methodology in Psychology

Replication and Extension of Aiken, West, Sechrest, and Reno's (1990) Survey of PhD Programs in North America

Leona S. Aiken, Stephen G. West, and Roger E. Millsap
Arizona State University

In a survey of all PhD programs in psychology in the United States and Canada, the authors documented the quantitative methodology curriculum (statistics, measurement, and research design) to examine the extent to which innovations in quantitative methodology have diffused into the training of PhDs in psychology. In all, 201 psychology PhD programs (86%) participated. This survey replicated and extended a previous survey (L. S. Aiken, S. G. West, L. B. Sechrest, & R. R. Reno, 1990), permitting examination of curriculum development. Most training supported laboratory and not field research. The median of 1.6 years of training in statistics and measurement was mainly devoted to the modally 1-year introductory statistics course, leaving little room for advanced study. Curricular enhancements were noted in statistics and to a minor degree in measurement. Additional coverage of both fundamental and innovative quantitative methodology is needed. The research design curriculum has largely stagnated, a cause for great concern. Elite programs showed no overall advantage in quantitative training. Forces that support curricular innovation are characterized. Human capital challenges to quantitative training, including recruiting and supporting young quantitative faculty, are discussed. Steps must be taken to bring innovations in quantitative methodology into the curriculum of PhD programs in psychology.

Keywords: quantitative curriculum, training, statistics, research design, measurement

Quantitative methodology, broadly defined, occupies a unique and ubiquitous role in the PhD curriculum in psychology. The introductory statistics sequence in the first year of graduate training is the last bastion of a core curriculum in psychology. Courses in research design and measurement draw students from across the full range of subdisciplines of psychology. Indeed, common training in quantitative methodology may be the one aspect of doctoral education that continues to unify the discipline of psychology. Given its central position, a careful examination and evaluation of quantitative training in the PhD curriculum across the broad discipline of psychology is warranted. Our central foci are the degree to which this training reflects the advances in methodology over the decade from 1990 to 2000 and the degree to which this training supports the research and application endeavors of psychology in the 21st century. In this article, we document the quantitative curriculum at the end of the past decade, on the basis of a survey of all PhD programs in psychology in North America. We also characterize the evolution of the quantitative curriculum over the preceding 13 years, on the basis of our previous survey of the same population of programs (Aiken, West, Sechrest, & Reno, 1990).

This is an exciting and even exhilarating time for the development of new quantitative methodologies. We include in quantitative methodology three broad topics: statistics, measurement (test theory, test construction, and/or other measurement techniques, e.g., scaling), and research design (the structure of experimental, quasi-experimental, and observational studies). The great strides in quantitative methodology are closely aligned with the increasing diversity and complexity of research questions that are now becoming central in different areas of psychology. To cite but a few examples, in statistics, the development of growth curve modeling (Bollen & Curran, 2006; Singer & Willett, 2003) provides characterizations of trajectories of gain and decline over time. This development goes hand in hand with an increasing emphasis on the study of life span development, both normal and pathological, both natural and following interventions. Multilevel statistical models (Hox, 2002; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999) permit researchers to simultaneously study

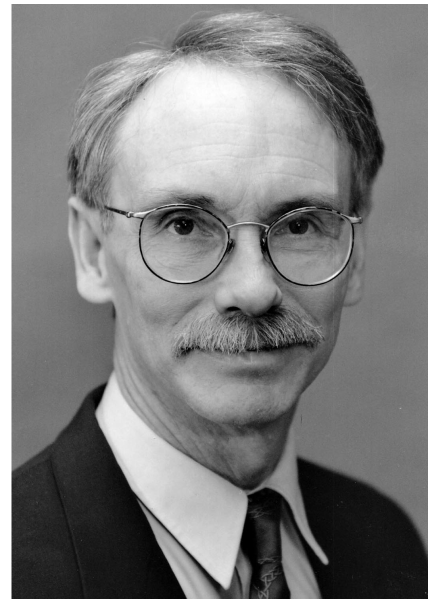
Leona S. Aiken, Stephen G. West, and Roger E. Millsap, Department of Psychology, Arizona State University.

We thank the College of Liberal Arts and Sciences, Arizona State University, and the College Board, New York, for their support of this project. We thank Robert Cudeck, Abigail Panter, William Shadish, Patrick Shrout, and Greg Stone for their review of early versions of the questionnaire. We thank Jennifer Anderson for her assistance in the open-ended coding. We are indebted to the participants who meticulously characterized the methodology curricula of their departments in our extensive survey instrument.

Correspondence concerning this article should be addressed to Leona S. Aiken, Department of Psychology, Arizona State University, Tempe, AZ 85287-1104. E-mail: leona.aiken@asu.edu

Vol. 63, No. 1, 32–50 DOI: 10.1037/0003-066X.63.1.32

Leona S. Aiken



Stephen G. West

multiple influences on psychological outcomes that occur at different levels of analysis: the individual, the dyad, families, small groups, or even large aggregations, such as schools and communities. Natural dependencies in data (clustering) have now become a feature to be understood and exploited rather than avoided. In measurement, the current study of measurement invariance (the extent to which measures have the same quantitative meaning across groups; Embretson & Reise, 2000; Millsap & Meredith, 2007; Widaman & Reise, 1997) is fundamental to the new emphases on the study of diversity across gender, ethnic, and language groups in psychology and on the renewed interest in cross-cultural research.

In research design, new methods have been developed for studying mediation (MacKinnon, 2008), for providing proper estimates of treatment effects even when participants choose not to receive the treatment (Angrist, Imbens, & Rubin, 1996; West & Sagarin, 2000), and for equating groups when participants cannot be randomized to treatment and control groups (Rosenbaum & Rubin, 1983; West & Thoemmes, in press). These new developments reach the psychology community directly through accessible texts (e.g., Embretson & Reise, 2000; Singer & Willett, 2003), through readily available software (e.g., SAS, SPSS), and through articles focused on software-based implementation of new statistical methods (e.g., Peugh & Enders, 2005; Singer, 1998). Begun in 1996, the American Psychological Association (APA) journal *Psychological Methods* has had the mission of making many of these new developments more available to both methodological experts and practicing researchers.

Our first purpose here is to examine the diffusion of innovation in quantitative methodology into the PhD curriculum from approximately 1990 to 2000. To this end, we provide an in-depth characterization of the doctoral-level quantitative methodology curriculum, which is based on a survey conducted in the late 1990s of over 200 PhD programs in North America. Our work is a replication and extension of Aiken et al.'s (1990) survey of the same population of doctoral programs in the late 1980s. Following in the footsteps of Aiken et al. (1990), we examine the content of the PhD curriculum in statistics, measurement, and research design, document requirements for quantitative methodology training by substantive specialty, and characterize judgments by informed faculty of the methodological competency of new PhDs from their programs. We also used reputation rankings to compare the methodolog-

Roger E. Millsap

ical training of elite (top 25) programs with the training of other programs (Goldberger, Maher, & Flatteau 1995). We re-examine the current status of five conclusions that emerged from Aiken et al.'s (1990) study:

1. Current PhD students are receiving traditional training in methodology and statistics, training that primarily supports laboratory rather than field research . . . .
2. Even with "ideal training" in a first-year graduate sequence, supplementary training is required . . . .
3. Measurement has declined substantially in the curriculum . . . .
4. Training in new techniques and methodologies is generally unavailable within the psychology curriculum. . ..
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. (p. 730)

Our replication gains important benefits from our previous baseline survey: We can now directly examine the changes in the curriculum in virtually the entire population of PhD programs in North America, focusing on the response of psychology departments to important issues that surfaced in our original article.

In the present survey, we also revisit the question of the adequacy of psychology's human capital to deliver training in new quantitative methods, an issue that was identified in our original survey. Even a cursory review of job advertisements for those with PhDs in psychology reveals a large number of available positions in quantitative methodology. There is a steady stream of openings for quantitative positions at major research universities, in teaching-oriented academic institutions, in research laboratories in universities, federal and state agencies, and the private sector, as well as in the testing industry. The demand seems ever expanding, and it outruns supply (Clay, 2005; Herszenhorn, 2006).

Method

Briefly, in April 1998, we mailed questionnaires asking about quantitative curriculum and quantitative personnel to all 234 departments identified in *Graduate Study in Psychology, 1997* (American Psychological Association, 1997) as conferring the PhD in at least one area of psychology. Of these departments, 86% ($n = 201$) responded. Of these 201 departments, 22, or 88%, of the 25 elite departments responded (Goldberger et al., 1995). Specific respondents included faculty responsible for teaching quantitative courses (86%); the remainder were mainly program administrators, some of whom may also have had specific quantitative expertise. A more detailed account of our methods and a description of the questionnaire are provided in the Appendix.

Data Reporting

Throughout the narrative, we use the terms *department* and *program* interchangeably; we use *area* or *concentration* to refer to a particular substantive specialty (e.g., clinical). We summarize our outcomes in a series of tables that provide our findings for the present 1998–1999 survey. As appropriate, we also report data from the survey by Aiken et al. (1990) to facilitate comparisons over time. Data reported in Aiken et al. (1990) were gathered in 1986; thus, tables refer to the 1986 survey.

Results

Results are organized into three broad sections: (a) the quantitative presence in psychology departments, including program characteristics, the quantitative curriculum, the competency in quantitative methods of those with PhDs, (b) evolution of the quantitative curriculum over time between surveys, and (c) support for quantitative training, including personnel and resources. Throughout the narrative, we highlight what we believe to be the most striking findings; we hope that readers will review the far richer information in the accompanying tables.

The Quantitative Presence in Psychology Departments

Program Characteristics

Table 1 characterizes the programs and contrasts the full complement of programs ($n = 201$) with the subset of elite institutions ($n = 22$). In all, 15% of programs (36% of elite programs) had a quantitative area (e.g., applied statistics, psychometrics, research methods, and mathematical psychology). Beyond concentrations leading to a PhD in quantitative methodology per se, fully 49% of programs reported that students in substantive areas could obtain a minor in quantitative methods (whether formal or informal) by taking a number of quantitative courses (68% of elite departments, 89% of departments with quantitative concentrations).

The Quantitative Curriculum in Psychology

The questionnaire listed a broad range of traditional and new topics that might be part of a quantitative methodology curriculum. Table 2 documents regular coverage of these topics, defined as coverage at least every two years. Duration of coverage is characterized in three categories, of which the first two overlap: (a) full course (i.e., a complete semester, a complete trimester, or a complete quarter), (b) at least a partial course (i.e., half of a semester, half of a trimester, or a complete quarter), or (c) none (i.e., no coverage in the curriculum). Finally, for each topic, we documented the availability of classes on campus outside the psychology department.

Statistics. The present survey shows that the old standards of psychology (e.g., analysis of variance [ANOVA]) were regularly taught in psychology. Structural equation modeling (SEM) emerged as a topic regularly taught in half of all departments. Slightly fewer than half of all departments offered a full SEM course. Coverage of specialized (and newer) statistical methods (e.g., multilevel modeling) was more sparse; when offered, treatment was limited to a brief segment in the curriculum.

Table 1
Program Demographics

Program characteristic	All programs (n = 201)	Elite programs (n = 22)
First-year class size (median)	13.6	14.5
No. of full-time faculty (median)	25.2	32.5
No. of full-time faculty who teach statistics, measurement, or general research design (median)	3.4	3.8
No. of faculty teaching statistics who were trained in statistics or research methods (median)	0.5	0.9
Schools with at least one faculty member trained in statistics or research methods	49%	45%
No. of faculty teaching exclusively quantitative methods (median)	0.7	0.8
Program offers PhD in quantitative area	15%	36%
Curriculum includes introductory graduate statistics course	100%	100%
Introductory course is offered in the department	91%	91%
Length of introductory course (% one academic year in length)	70%	62%
Course is required of all PhD students	97%	100%
Department regularly has students take statistics or measurement courses outside the Department of Psychology	36%	54%

Measurement. In all, 64% of all departments (50% of elite departments) provided a PhD-level course in measurement (including test theory, test construction, and/or other measurement techniques, e.g., scaling). Given the concerns about measurement training raised in our prior survey (see also Lambert, 1991; Meier, 1993; Merenda, 1996), we separately assessed training in classical test theory, in item response theory (IRT), and in test construction (see Table 2). The low rate of regular coverage of classical test theory, IRT, and even test construction is disconcerting. Moreover, coverage was typically brief. In all, 27% of programs combined short segments of classical test theory, IRT, and test construction in a single measurement course. Elite departments were less likely to provide regular coverage of most measurement topics than the general population of departments.

Research design. Almost all departments presented some general coverage of research design at least every two years. More specialized topics in research design, particularly those related to field research (e.g., quasi-experimentation), were taught far less frequently and in briefer course segments. In all, 57% of departments (32% of elite) offered a general research design course (defined as covering the structure of experimental, quasi-experimental, and/or observational studies); on average, 85% of students in those departments (61% of elite) took the course. Specific areas also offered their own research methods courses (see Table 3). However, 27% of substantive area units did not teach their own research methods course and also had no departmental research design course on which to draw.

Computer applications. Computer applications were part of the regular curriculum in the majority of departments (see Table 2).[1] Not unexpectedly, most instruction was in standard statistical packages (e.g., SPSS, SAS). Specialized application software was often taught in advanced courses, for example, Mplus (L. K. Muthén & Muthén, 1998–2007) and BILOG (Zimowski, Muraki, Mislevy, & Bock, 2006). In all, 72% of departments (82% of elite) had a PC-based computing laboratory for quantitative instruction; 92% (95% of elite) had such labs on campus.

More mathematical areas. More mathematical areas, including the mathematical foundations of statistics, mathematical psychology, and nonlinear modeling, were rare in the curriculum, except for mathematical psychology in elite programs (see Table 2).

Campus resources for quantitative training. More than half of the departments could draw on resources elsewhere on campus for training that was appropriate in topic and level for their students (see Table 2).

Introductory Statistics Sequence: Offering, Requirement, Content

Offering and requirement. The introductory graduate statistics course was universally a part of the PhD curriculum (see Table 1), offered yearly by 96% of programs. In 30% of departments (38% of elite), the course was shorter than one year.

Content of the introductory sequence. Respondents indicated whether each of the topics listed in Table 4 was covered in the introductory course sequence (a) in depth, to the point that students could perform the analysis in question themselves, (b) as an introduction to the topic to acquaint students with concepts, or (c) not at all. Most programs provided in depth coverage of ANOVA and multiple regression (MR); MR occupied a median of 7.3 weeks in the introductory course (9 weeks in elite departments). Coverage of some core topics in MR was weak (e.g., regression diag-

[1] Fully 94% of all programs (95% of elite programs) taught standard statistical packages (e.g., SPSS, SAS) in the introductory statistics core courses; this was so in 82% (73% elite) of advanced quantitative courses. Modern interactive object-oriented software for graphically oriented computing—for example, ARC (R. D. Cook & Weisberg, 2004), DataDesk (Vellman, 2006), R (R Project for Statistical Computing, 2007), or S-PLUS (Insightful Corporation, 2007)—was used less frequently, in 11% (5% elite) of introductory quantitative courses, and 25% (50% elite) of advanced quantitative courses. Specialized software for particular applications—for example, EQS, AMOS, LISREL, Mplus for SEM; HLM, MLwiN (Rasbash, Steele, Browne, & Prosser, 2005) for hierarchical linear modeling; and BILOG for measurement applications—was used in 74% of advanced courses (72% in elite programs).

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

	Topic taught at least every two years (%)		Duration of coverage (%)			Available on campus (%)		Duration of coverage: 1986 study (%)		
Curriculum	All	Elite[a]	Full[b]	Partial[c]	None[d]	All	Elite[a]	Full[b]	Partial[c]	None[d]
Statistics core										
Analysis of variance	95	91	71	83	3	79	73	65	88	3
Multiple regression	95	91	56	76	3	81	77	36	68	8
Multivariate analysis	80	76	55	67	9	76	77	48	63	14
Factor analysis	74	73	18	36	11	62	50	20	36	20
Structural equation modeling[e]	52	46	42	47	27	59	54	14	18	45
Specialized statistics content										
Multilevel models	34	52	11	16	48	40	41	—	—	—
Longitudinal data analysis[f]	30	29	13	18	50	55	68	4	6	63
Meta-analysis	41	33	19	21	40	32	31	—	—	—
Categorical data analysis	43	52	19	22	37	59	64	—	—	—
Modern missing data treatment	24	25	3	3	66	27	36	—	—	—
Nonparametrics	50	30	17	19	36	61	59	—	—	—
Measurement and scaling										
Classical test theory	64	57	22	35	20	49	36	—	—	—
Item response theory	40	29	9	13	42	38	36	—	—	—
Classical and/or item response theory[g]	64	57	24	36	21	65	45	31	45	30
Test construction	61	43	20	29	24	53	46	13	25	40
Multidimensional scaling[h]	25	52	12	16	50	37	46	16	21	38
Research design/methods										
Research design[i]	92	80	72	80	4	66	54	56	70	13
Quasi-experimental design	66	40	18	27	20	54	24	14	28	23
Survey research	36	25	17	21	51	62	64	10	15	48
Program evaluation[j]	33	5	31	36	7	50	25	29	36	37
Survey sampling	24	15	8	11	63	59	64	—	—	—
Other topics										
Computer applications	66	63	30	39	24	62	55	41	53	13
Mathematical foundations of statistics	29	25	10	12	67	61	68	—	—	—
Mathematical psychology	9	38	20	21	75	18	18	22	24	64
Nonlinear modeling	21	16	4	7	62	44	36	—	—	—
Epidemiology	4	5	4	5	90	36	59	—	—	—

Note. The first seven columns of numbers refer to the current survey; the final three columns refer to the 1986 survey reported in Aiken et al. (1990). Values are the percentages of schools responding affirmatively to the topic. Dashes indicate that a topic was not addressed in the 1986 survey.
[a] Values under the "Elite" headings are for the 22 elite universities. [b] A full course was included in the curriculum (i.e., a complete semester, complete trimester, or complete quarter). [c] At least half a semester of coverage (including a half semester, a half trimester, or a complete quarter) was included in the curriculum. [d] Topic was not included in the Department of Psychology curriculum. [e] This was called *causal modeling* in the previous survey. [f] This was called *time series* in the previous survey. [g] This was called *test theory* in the previous survey. [h] This was called *scaling* in the previous survey. [i] This was called *research methods* in the previous survey. [j] This was called *evaluation research* in the previous survey.

nostics, interactions). Only one third of programs provided in-depth training in statistical power analysis.[2]

Course Offerings and the Extent of the Methodology Requirement

A comparison of Tables 2 and 4 (i.e., the full statistics curriculum vs. coverage in the introductory course [sequence]) makes it clear that much content was reserved for more advanced courses, particularly multivariate analysis, SEM, longitudinal methods, and missing data. The median number of quantitative courses (statistics plus measurement plus design) taught regularly during the year beyond the introductory course sequence was 2.8 courses, but there

[2] Respondents listed other topic areas included in the introductory sequence: 4% of programs mentioned factor analysis; 4% mentioned topics in multivariate analysis; 4% mentioned specific topics in regression analysis (e.g., redundancy analysis); 3% mentioned categorical data analysis (e.g., loglinear models); 2% mentioned Bayesian statistics; and 2% mentioned robust estimates.

Table 3
Departmental Offerings in Research Design and Area Offerings in Research Methods

Area	No. of programs with area	Area offers own research methods course (%)[a]	Department offers research design course (%)[b]	Neither departmental nor area methods course (%)[c]	1986 survey: Neither departmental nor area methods course (%)[c]
Clinical	146	61	24	15	16
Counseling	17	47	29	24	23
Developmental	107	34	39	27	28
Cognitive	122	25	38	37	36[d]
Biopsychology	110	22	43	35	—
Personality	49	27	40	33	49
Quantitative	29	32	25	43	—
Social	121	50	26	24	23
IO, engineering, human factors	61	55	35	10	16[e]

Note. The first four columns of numbers refer to the current survey; the final column refers to the 1986 survey reported in Aiken et al. (1990). Dashes indicate that research design offerings were not considered for biopsychology and quantitative psychology in 1986. IO = industrial–organizational psychology.
[a] Percentage of all programs in an area that offer their own research methods course. [b] Percentage of all programs in an area whose departments offer a general research design course. [c] Percentage of all programs in an area that do not offer their own research methods course and also do not have a departmental course in research design. [d] Area was identified as *experimental* in the 1986 survey. [e] Area was identified as *applied* in the 1986 survey.

was great variation. Only 17% of programs reported no such courses. In all, 16%, 12%, 20%, 9%, and 9% of programs reported one, two, three, four, or five such additional courses, respectively; the remaining 18% of programs reported more than five such courses.

Table 5 shows the extent of the requirements in statistics and measurement, as well as the extent of the quantitative course work actually taken, stratified by substantive area. The statistics requirement was essentially universal, whereas the measurement requirement varied widely across areas. With the exception of quantitative concentrations, most areas required just over a year of statistics. The measurement requirement was brief, with a median of 0.15 years or about 4.5 weeks across all areas. Median years of course work actually taken hovered at about 1.6. Given that the introductory sequence was 1 year in length in 70% of the programs, the typical PhD student took slightly over half a year more of statistics and measurement beyond the first year sequence.

Competency of New PhDs to Utilize Quantitative Methodology

Participants judged the competencies of their recent PhD graduates to utilize quantitative methodology. Specifically, participants rated whether most or all (>75%), some (25%–75%), few (1%–25%), or none (0%) of their students would be well enough acquainted with a variety of methods to apply them in their own research (see Tables 6, 7, and 8).

Statistics. Table 6 indicates that competencies were judged to be high in traditional topics. Yet, even within the broad classic topics of ANOVA and MR, there were noteworthy areas of limited competency (e.g., use of regression diagnostics that are critical for assessing the robustness of regression models). About half of all programs judged that few or none of their students were able to apply SEM. Most programs judged that few or none of their students could apply methods of multilevel data and data over time or could apply modern methods of missing data, or logistic regression, which are increasingly the state of the art for binary diagnostic outcomes in clinical or health psychology.

Measurement. An analysis of competencies in measurement raises grave concern about the most fundamental issues for adequate measurement in psychological research (see Table 7). Fewer than half of the respondents judged that most of their students could assess the reliability of their measures; only one fourth of respondents judged that most of their students could utilize methods of validity assessment. Only one fourth of respondents indicated that most of their students were competent at test construction or item analysis.

Research design. Judged competency was high in the design of laboratory experiments but was much lower in the design of field experiments (see Table 8). Competency was judged to be almost nonexistent in quasi-experimental designs and in designs involving data collected over time.

Do Elite Departments Provide Better Training in Methodology?

One possible hypothesis for the overall weak showing in competency in quantitative methodology is that such training was, in fact, stronger in elite departments but that increasing numbers of students were being trained in other than elite PhD programs. We also considered but failed to support this hypothesis in Aiken et al. (1990). In fact, the present survey

Table 4
Contents of the Introductory Statistics Sequence

Introductory statistics content	In-depth coverage (%)[a]		No coverage (%)		1986 survey (%)[d]	
	All[b]	Elite[c]	All[b]	Elite[c]	In-depth[a]	No coverage
Data description						
Traditional data description	44	27	7	5	—	—
Modern graphical displays	10	14	56	40	—	—
Analysis of variance (ANOVA)						
Multifactor ANOVA	80	87	3	0	73	6
A priori comparisons	79	77	0	0	69	5
Post hoc comparisons	79	77	0	0		
Repeated measures handled by traditional factorial ANOVA	72	56	4	4	73	7
Analysis of covariance	52	50	7	0	39	8
Incomplete designs	13	27	37	23	11	33
Regression						
Multiple regression	78	77	2	5	63	8
Hierarchical regression (sets of predictors)	57	64	14	14		
ANOVA as special case of regression	52	54	7	5	38	14
Interactions in multiple regression	42	41	17	5	—	—
Regression diagnostics	31	46	18	9	—	—
Regression graphics	25	32	21	9	—	—
Logistic regression	9	18	55	59	—	—
More advanced approaches						
Repeated measures handled by multivariate procedures	29	29	25	19	21	37
Multivariate procedures (MANOVA, canonical, discriminant)	26	14	37	46	21	40
Structural equation modeling	12	0	55	68	5	59
Other areas						
Power analysis	36	41	5	9	18	27
Computer intensive statistics	4	9	56	48	—	—
Significance testing debate	32	29	14	14	—	—

Note. The first four columns of numbers refer to the current survey; the final two columns refer to the 1986 survey reported in Aiken et al. (1990). Dashes indicate that a topic was not addressed in the 1986 survey. MANOVA = multivariate analysis of variance.
[a] Coverage was included to the extent that students could use the technique in their own research. [b] All $N = 201$ programs in the current survey. [c] The $n = 22$ elite programs in the current survey, a subset of all 201 programs. [d] All $N = 186$ programs in the original survey reported in Aiken et al. (1990).

showed that there was less formal instruction in measurement and research design in elite programs but somewhat more coverage in a few topics, including mathematical psychology. More students in elite schools took statistics classes outside psychology. Judged competencies in statistics were higher in elite programs than overall (Table 6) but were lower in elite programs than overall in both measurement (Table 7) and research design (Table 8).

Evolution of the Curriculum Over Time

A comparison of the curriculum in our original survey with that of the current survey identified several areas in which there were gains in quantitative methodology offerings (see Table 2 for the overall curriculum and Table 4 for the introductory doctoral sequence).[3]

These gains were accompanied by parallel gains in the judged competence of those with new doctoral degrees to use quantitative methodologies in these areas (see Tables 6, 7, and 8). The gains we observed could be manifested in two distinct ways: (a) an increase in the percentage of programs reporting in-depth coverage or lengthy coverage or (b) a decrease in the percentage of programs offering no coverage at all.

Extent of Requirements and Offerings in Quantitative Methodology

Taken across areas of psychology, the combined median number of required years of statistics plus measurement was 1.2, the same as in our original survey (see Table 5).

[3] We caution that some topics included in the present survey were not included in the previous survey and that we used somewhat different terminology for a few topics in our two surveys, as noted in footnotes e through j of Table 2.

Table 5
Total Requirements in Statistics and Measurement

Area	Requires at least one statistics course (%)	Requires at least one measurement course (%)	Mean no. of years of required statistics courses	Mean no. of years of required measurement courses	Mean no. of years of statistics and measurement courses required	Mean no. of years of statistics and measurement courses taken	1986 survey: Mean no. of years of statistics and measurement courses required
Clinical	99	51	1.1	0.26	1.4	1.5	1.2
Counseling	93	64	1.1	0.36	1.5	1.6	1.2
Developmental	99	13	1.1	0.08	1.2	1.4	1.2
Cognitive	98	11	1.1	0.07	1.1	1.3	1.2
Biopsychology	98	11	1.0	0.06	1.1	1.1	—
Personality	98	28	1.0	0.15	1.2	1.6	1.1
Quantitative	97	42	1.9	0.22	2.2	3.1	2.0
Social	99	16	1.1	0.09	1.2	1.6	1.2
IO, engineering, human factors	100	52	1.3	0.31	1.6	2.0	1.4

Note. The first six columns of numbers refer to the current survey; the final column refers to the 1986 survey reported in Aiken et al. (1990). One semester and one trimester were counted as 0.50 year; one quarter was counted as 0.33 years. IO = industrial–organizational. A dash indicates that a topic was not addressed in the 1986 survey.

The number of quantitative courses that programs offered beyond the introductory course remained stable between surveys in 62% of programs, increased in 28% of programs, and declined in the remaining 10%.

Statistics Curriculum

As revealed in Table 2, two statistical areas showed appreciable increases in coverage between surveys: MR and SEM. For both areas, the number of departments offering full courses increased. For SEM, the number of departments with no coverage declined as well. Judged competencies of new PhDs in statistical analysis mirrored these curricular innovations. More programs judged that most or all of their students could apply ordinary least squares regression in their own research (58% to 85%). Fewer programs judged that few or none of their students could use SEM in their own research (81% to 50%).

Measurement Curriculum

The measurement curriculum appears to have changed with the introduction of short segments into the curriculum (i.e., a decline in the percentage of programs offering no coverage in test theory, test construction, and multidimensional scaling; see Table 2). Judged competencies in measurement topics increased slightly over time as well (see Table 8), with the largest increase in reliability assessment. For newer measurement topics, IRT and generalizability theory, there were minor declines in the percentage of programs indicating that few or none of their graduates could utilize these techniques. Regrettably, the proportion of programs that indicated that most or all of their PhDs could apply classical or more recent measurement approaches still remained low.

Research Design

Table 2 reveals a slight increase in coverage of research design in a general departmental course (57% of programs, up from 49% in our original survey). However, the percentage of program areas with no access to either a departmental research design course or area-specific research methods remained essentially constant between surveys (see Table 3). Yet, we observed decreases in judged competency between our two surveys—in more specialized designs (Person × Situation designs and single subject designs). For many topics in research design, judged competency remained at an extraordinarily low level over time (see Table 8). Of the three broad areas of statistics, measurement, and research design, we observed the least gain and the weakest state of affairs in design.

The Introductory Course Sequence

The percentage of departments whose introductory statistics sequence was shorter than 1 year rose from 23% to 30%. MR and related topics enjoyed more in-depth coverage (see Table 4). In-depth coverage of statistical power analysis doubled, with a noteworthy decrease in introductory courses containing no coverage at all.

Fine-Grained Assessment of Curricular Innovation

Participants listed, in open-ended fashion, up to three important topics added to their departments' quantitative curriculum during the past decade (see Table 9). The three most frequently mentioned were SEM, topics in measurement, and meta-analysis. Topics that were being newly developed during the 1990s were infrequently added (e.g., multilevel modeling, longitudinal data analysis). Depart-

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

	Programs indicating whether graduates can apply techniques to their own research (%)					
	Most or all[a] (≥75%)		Few or none[b] (≤25%)		1986 survey[e]	
Statistics technique	All[c]	Elite[d]	All[c]	Elite[d]	Most or all[a]	Few or none[b]
Data description						
Traditional data description (e.g., skew, histograms)	89	100	2	0	—	—
Modern graphical data display (e.g., quantile–quantile [QQ] plots, kernel density estimates)	6	14	76	62	15	55
Analysis of variance (ANOVA) and related topics						
Multifactor ANOVA	80	100	6	0	81	4
A priori contrasts/focused contrasts	75	100	4	0	76[f]	6[f]
Post hoc comparison procedures	79	100	2	0		
Repeated measures by factorial ANOVA	71	86	9	5	74	5
Repeated measures handled by multivariate procedures	33	38	28	24	22	37
Analysis of covariance and alternatives	50	57	13	0	38	22
Incomplete designs (e.g., Latin squares)	14	24	58	29	—	—
Multiple regression and related topics						
Ordinary least squares multiple regression	85	95	4	0	58	6
Continuous variable interactions in regression	42	52	22	0	—	—
Logistic regression	9	19	59	43	—	—
Regression diagnostics	23	33	45	29	8	68
Nonlinear models	7	4	77	71	—	—
Multivariate analysis and related topics						
Matrix algebra for multivariate analysis	12	14	57	57	—	—
MANOVA/discriminant analysis	34	33	25	38	18	34
Structural equation modeling						
Path analysis	12	24	46	48	—	—
Structural equations with latent variables	10	19	50	52	2	81
Methods for clustered and over-time data						
Multilevel (random coefficient, hierarchical) models	5	5	74	62	—	—
Time series analysis	3	14	84	87	1	86
Longitudinal data analysis (e.g., survival, growth modeling)	2	0	81	67	—	—
Estimation approaches and missing data						
Robust statistics and robust estimation	6	10	77	67	—	—
Bayesian and empirical Bayes methods	2	14	89	86	—	—
Computer intensive statistics (bootstrapping, resampling)	4	10	81	67	—	—
Modern treatment of missing data (e.g., expectation maximization [EM] algorithm)	3	0	85	86	—	—
Treatment of categorical data						
Basic nonparametric procedures	40	43	30	29	48	20
Categorical data analysis	20	29	48	29	—	—
Meta-analysis	7	5	60	57	—	—
Statistical computing: PC-based statistical packages	82	100	4	0	—	—

Note. The first four columns of numbers refer to the current survey; the final two columns refer to the 1986 survey reported in Aiken et al. (1990). Dashes indicate that a topic was not addressed in the 1986 survey. MANOVA = multivariate analysis of variance.
[a] Most or all recent graduates (≥75%) can apply the technique to their own research. [b] Few or none of recent graduates (≤25%) can apply the technique to their own research. [c] All $N = 201$ programs in the current survey. [d] The $n = 22$ elite programs in the current survey, a subset of all 201 programs. [e] All $N = 186$ programs in the original survey reported in Aiken et al. (1990). [f] The 1986 survey combined contrasts and comparisons in a single item.

ments with quantitative PhD programs were much more likely to have added new topics in multilevel modeling (20%), in longitudinal data analysis (23%), and in cognitive-neuroscience related topics (17%).

What Is Lacking in the Curriculum?

Respondents listed, in open-ended fashion, up to three important topics that were not available anywhere on campus at a level appropriate for psychology PhD students (see

Table 7
Judged Competencies of Graduates to Apply a Variety of Measurement Approaches in Their Own Research

Measurement approach	Programs indicating whether graduates can apply measurement approaches in own research (%)					
	Most or all[a] (≥75%)		Few or none[b] (≤25%)		1986 survey[e]	
	All[c]	Elite[d]	All[c]	Elite[d]	Most or all[a]	Few or none[b]
Unidimensional scaling	19	14	52	48	5	69
Multidimensional scaling	10	14	66	48	2	74
Classical test theory	30	14	36	33	19	53
Item response theory	8	5	60	71	6	76
Item analysis	22	10	37	50	17	57
Reliability assessment	46	29	19	24	27	38
Test construction	26	10	33	50	—	—
Generalizability theory	9	0	64	90	6	75
Use of tests in selection	15	0	49	85	3	80
Evaluation of test bias	14	0	53	90	1	89
Methods of validity assessment	28	0	31	55	22	44

Note. The first four columns of numbers refer to the current survey; the final two columns refer to the 1986 survey reported in Aiken et al. (1990). Dashes indicate that a topic was not addressed in the 1986 survey.
[a] Most or all recent graduates (≥75%) can apply the technique to their own research. [b] Few or none of recent graduates (≤25%) can apply the technique to their own research. [c] All N = 201 programs in the current survey. [d] The n = 22 elite programs in the current survey, a subset of all 201 programs. [e] All N = 186 programs in the original survey reported in Aiken et al. (1990).

Table 9). The results largely mirror our list of topics added in some departments, indicating that a separate set of psychology departments perceived similar needs but did not add these same topics to the curriculum. Innovation in quantitative methods diffused differentially into programs between surveys. Of the three broad areas of statistics, measurement, and research design, additional topics in research design were by far the least likely to be added to the curriculum (1.5%) or to be perceived as needed (0.5%).

Table 8
Judged Competencies of Graduates to Apply a Variety of Research Designs in Their Own Research

Research design	Programs indicating whether graduates can apply designs in own research (%)					
	Most or all[a] (≥75%)		Few or none[b] (≤25%)		1986 survey[e]	
	All[c]	Elite[d]	All[c]	Elite[d]	Most or all[a]	Few or none[b]
Design of laboratory experiments	81	85	4	5	83	4
Design of field experiments (basic research in field settings)	46	30	11	30	42	13
Program evaluation	10	0	55	80	—	—
Experimental personality designs (Person × Situation designs)	18	5	47	50	25	34
Time series designs	4	0	72	85	4	79
Regression discontinuity designs	3	0	78	70	4	82
Nonequivalent control group designs	17	5	46	45	12	58
Longitudinal designs	11	5	44	50	12	52
Qualitative methodologies	8	0	67	85	5	68
Single subject designs	8	0	69	85	10	59

Note. The first four columns of numbers refer to the current survey; the final two columns refer to the 1986 survey reported in Aiken et al. (1990). Dashes indicate that a topic was not addressed in the 1986 survey.
[a] Most or all recent graduates (≥75%) can apply the technique to their own research. [b] Few or none of recent graduates (≤25%) can apply the technique to their own research. [c] All N = 201 programs in the current survey. [d] The n = 22 elite programs in the current survey, a subset of all 201 programs. [e] All N = 186 programs in the original survey reported in Aiken et al. (1990).

Table 9
Topical Areas Added to the Quantitative Curriculum in Past Decade and Topical Areas Lacking in the Quantitative Curriculum

Topic (topical area)	Programs (%) adding topical area (*n* = 200)[a]	Programs (%) lacking topical area (*n* = 200)[b]
Topics in structural equation modeling (path analysis, causal modeling, covariance structure analysis, multivariate latent approach, confirmatory factor analysis, dynamic factor analysis)	44.5	14.0
Topics in measurement (psychometrics, scaling, test construction, generalizability theory, advanced tests and measurement, item response theory)	21.0	20.5
Meta-analysis	21.0	10.5
Topics in classic multivariate analysis (canonical, cluster, exploratory factor analysis, multivariate analysis of variance, multivariate statistics)	12.0	5.0
Topics in categorical data analysis (multivariate categorical analysis, logistic/logit/probit, loglinear, nonparametric models)	12.0	10.0
Topics in multilevel modeling (multilevel modeling, random effects regression)	11.5	17.0
Topics in longitudinal data analysis (growth curve modeling, latent growth models, longitudinal growth modeling, trajectories of change, methodology for longitudinal data, survival analysis, time series analysis)	9.5	7.5
Multiple regression	8.5	0.5
Statistical power analysis	6.5	0.0
Topics related to cognitive psychology and neuroscience (mathematical psychology, artificial intelligence, behavior genetics, quantitative genetics, chaos data analysis, computational modeling, dynamical systems, neural networks, neuroscience models of stochastic processes)	5.0	4.0
Computer applications	4.0	0.5
Program evaluation	3.0	1.5
Missing data	2.5	8.0
Nonlinear modeling	1.5	4.5

[a] Additional infrequently added topics include effect sizes (2.5%), hypothesis testing (2.5%), robust statistics (2%), exploratory data analysis (1.5%), research design (1.5%), and graphics (1.5%). [b] Additional topics mentioned infrequently as missing from the curriculum include statistical power analysis, robust statistics, Bayesian methods, and survey research (2% each); bootstrap estimators (1.5%); sampling (1%); clinical decision making, judgment and decision making, data mining, econometric modeling, epidemiology, exact hypothesis testing, event history analysis, Gibbs sampler, mathematical foundations, philosophy of science foundations, programming experimental software, and research design (0.5% each).

Support for Quantitative Training

Quantitative Faculty and Quantitative Training Resources

Psychology faculty who provide quantitative training. An average of 16% of all psychology department faculty taught statistics, research design, and/or measurement (excluding clinical assessment). We divided these faculty into three groups. Group 1 included faculty trained as quantitative methodologists who identified themselves primarily as quantitative specialists (20% of all faculty teaching quantitative methods, 28% in elite programs). Group 1 tended to be clustered in quantitative psychology PhD concentrations. Group 2 included faculty trained in a substantive area with which they identified (42% of all faculty teaching quantitative methods, 31% in elite programs). Group 3 included faculty who were jointly trained in a substantive area and in quantitative methods and who identified with both (38% of all faculty teaching quantitative methods, 41% in elite programs). Only half of all programs had at least one quantitative methodologist (Group 1) on the faculty. In contrast, 75% of programs had a member of Group 2; 76% of programs had a member of Group 3. In fact, 81% of faculty teaching quantitative methods also taught substantive courses (68% in elite programs).

Additional sources of course work. In all, 20% of departments (18% of elite departments) employed adjunct faculty to teach graduate statistics. In 36% of programs (54% of elite programs), students regularly took graduate statistics or measurement courses in other departments. To the extent that elite departments are located in stronger universities overall, greater opportunity for appropriate outside training may exist.

Support for Training in Quantitative Methods: Faculty and Students

Support was available allowing faculty in 43% of programs and students in 30% of programs to attend special methodology workshops and conferences. In 19% of programs, there was a regular methodology brown bag meeting or other public forum in the department that provided informal consulting and training in quantitative methods. In 40% of programs, some faculty in substantive areas audited quantitative courses to gain new training in quantitative methods.

Resources for Obtaining a PhD in Quantitative Psychology

In the present survey, 31 programs (15%) offered a PhD concentration in a quantitative area. Across these 31 programs, there was a total of 47 first-year students; 9 programs had no-first year students. This is a noteworthy decrease from the 108 first-year quantitative students in our original survey, over a 50% decline![4] In the late 1990s, the total number of students enrolled in all these programs in all years was 183; of the programs, 3 had no students at all; the median total number of students per program was 4.70. We asked whether programs had had a quantitative program at any time between our two surveys. In all, 41 said *yes,* with 5 of these programs initiated since our original survey. However, of the 41 programs that existed at some point during this period, only 26 were still functioning (i.e., were still accepting doctoral students) at the time of the current survey; the remainder had functionally closed. A study carried out in late 2006 showed evidence of continued program attrition. There were only 25 quantitative PhD programs in North America, with a total of 40 first-year students; 7 of the 25 programs had no first-year students (APA Task Force to Increase the Quantitative Pipeline, 2007).

Hiring and Replacement of Faculty Who Teach Quantitative Methods

The number of faculty who taught quantitative methods remained stable between our surveys. In the five years just preceding the current survey, 41% of programs lost at least one such faculty member, and 50% of programs hired at least one such faculty member. This flux led to an average net increase across programs of 0.1 faculty members teaching quantitative methods.

Discussion

Five Conclusions a Dozen Years Later?

Our conclusions show considerable stability, with some indications of mixed change over a dozen years.

1. We stated in 1990, "Training in methodology and statistics . . . primarily supports laboratory rather than field work" (Aiken et al., 1990, p. 730). This conclusion is certainly true for research design. Laboratory methods receive far more coverage than field methods, and our PhDs are judged to be far more competent to conduct research in laboratory than the field.

Increases in training in MR and SEM support observational field research. Yet, topics appropriate to laboratory studies predominated in the introductory statistics course sequence. The training in quantitative methodology documented in our current survey does not support research as it is carried out in the field—with specialized populations, longitudinal designs, variables of interest measured at multiple levels (e.g., the individual, family, school, community), and limitations on the ability to assign participants randomly to treatment conditions or to maintain random assignment once it has occurred.

2. As in 1990, "Even with 'ideal training' in a first-year graduate sequence, supplementary training is required" (Aiken et al., 1990, p. 730). The introductory statistics course sequence, less than one year in length in 30% of programs, covered mainly ANOVA and increasingly MR. It is at best foundational for what is to follow in quantitative training. Improvement is needed in coverage of current topics (more complete teaching of regression analysis, including interactions, regression diagnostics, statistical power analysis).

3. Perhaps our most dire conclusion in 1990 was that "measurement has declined substantially in the curriculum" (Aiken et al., 1990, p. 730; see also Lambert, 1991; Meier, 1993; Merenda, 1996, 2003, 2006). There were some improvements in the measurement curriculum—a decrease in the percentage of programs with no coverage of key topics in measurement, the addition of topics in measurement to the curriculum by over 20% of programs, and increases in the judged competency in measurement of new PhDs. Nonetheless, we find it deplorable that a dozen years later, the measurement requirement occupies a median of only 4.5 weeks in the PhD curriculum in psychology. A substantial fraction of programs offered no training in test theory or test construction; only 46% of programs judged that the bulk of their graduates could assess even the reliability of their own measures. We conclude that coverage in measurement remained inadequate a dozen years later and that most graduates lacked fundamental competency in measurement at the point of our current survey.

4. Our conclusion that "training in new techniques and methodologies is generally unavailable within the psychology curriculum" (Aiken et al., 1990, p. 730) remains true. In many instances, critical newer topics in methodology were included neither in the psychology curriculum nor elsewhere on campus. This was true for 42% of programs for SEM, 48% for longitudinal data analysis, and 23% for multilevel models and IRT. This same lack of available coverage prevailed for more basic topics as well (e.g., test construction).

5. It appears even more true than before that "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" (Aiken et al., 1990, p. 730). Nonresponse to questions of availability of training on campus rose from 20% to 25%. We attribute this nonresponse to failure of awareness; this is the only area in which missing data were more than minimal.

[4] In Aiken et al. (1990), we reported that there were 108 quantitative PhD students across all years. However, our review of our original data revealed that the value 108 referred to the number of first-year quantitative PhD students across all quantitative programs.

Research Design—A New Area of Great Concern

Alongside measurement, we also have profound concern regarding training in research design. This is the only area in which we found slippage between the original and current surveys. Our current canon of experimental and quasi-experimental designs was largely developed in the 1950s, 1960s, and 1970s. Today, nearly all of the notable new developments in field research design occur outside of psychology, often in statistics departments (e.g., Rosenbaum, 2002; Rubin, 2005), where they are less likely to be noticed or understood by psychologists. These include techniques for treatment of broken randomized experiments (Barnard, Du, Hill, & Rubin, 1998), treatment noncompliance in experiments (Angrist, Imbens, & Rubin, 1996), and propensity scores in observational studies (Rosenbaum & Rubin, 1983). With the exception of occasional chapters reviewing this work (e.g., Shadish, Luellen, & Clark, 2006; West & Sagarin, 2000), there is little to inform psychologists about these new developments.[5] Psychologists must reinvigorate the teaching of research design to our next generation of graduate students, to bring new developments burgeoning in other fields into the mainstream of psychology.

What Forces Support Curricular Innovation?

We believe that three factors combine to support diffusion of innovation in statistics and measurement into the PhD psychology curriculum. SEM provides a case study. The first factor is a reciprocal interplay between the availability of an innovative new quantitative methodology and the movement of psychologists into new research areas. As researchers grasp the possibilities for the new methodology to answer important innovative research questions and in turn are stimulated by the new methodology to ask still other new research questions, they gain excitement about utilizing the new methodology. Peter Bentler's (1980) chapter in the *Annual Review of Psychology* and David Kenny's (1979) text, *Correlation and Causality,* introduced psychologists to the possibilities of SEM with latent variables. In turn, substantive researchers can drive methodological development as they become users, implement new applications, and discover unforeseen limitations of existing approaches.

Second is the availability of texts at an accessible level for both faculty and doctoral students. No accessible textbook existed before 1987 that provided practical guidance on how to conduct modern SEM analyses in the latent variable framework. In short order, two accessible texts, Hayduk (1987) and Loehlin (1987) appeared, followed by a complete higher level reference work two years later (Bollen, 1989).

Third is the availability of user friendly software. Software was certainly available for SEM by the early 1980s; however, this software demanded a strong background in the mathematics underlying SEM for proper use. In 1985, the EQS software (Bentler, 1985) made model specification and interpretation of results much more accessible. SEM software continues to become increasingly user friendly, resulting in ever more widespread use. For SEM, all three factors were in place by the late 1980s, and the percentage of programs with a full SEM course rose from 14% to 42% in the subsequent dozen years.

Innovations in the First Decade of This Century

The three conditions for curricular innovation are now in progress for several developments in statistics and one in measurement. In statistics, multilevel modeling and longitudinal growth modeling support the widespread movement of psychology into complex multilevel and longitudinal designs in field settings. Bryk and Raudenbush's (1992) text, the first comprehensive treatment of multilevel modeling, was challenging and focused on educational research. Recent texts (Hox, 2002; Kreft & de Leeuw, 1998; Snijders & Bosker, 1999) provided more accessible presentations of multilevel modeling. Singer and Willett (2003) and Bollen and Curran (2006) supported longitudinal growth modeling by psychologists. Both SAS Proc Mixed (SAS Institute, 1992) and stand alone HLM software (e.g., Bryk, Raudenbush, Seltzer, & Congdon, 1988) for multilevel modeling appeared early. Ease of access to multilevel modeling expanded with inclusion of mixed modeling in SPSS 11.0 (Norusis, 2002; SPSS, 2001). User friendly software for longitudinal modeling in the latent variable framework—for example, AMOS (SPSS, 2007), Mplus, and EQS 6 (Bentler, 1995)—complemented this development. This combination of new statistical techniques that answer new substantive questions, accessible texts, and accessible software has led to the rapid growth of training in and use of multilevel and latent growth curve models that psychologists are witnessing at present.

New methods for treating missing data (i.e., multiple imputation and full information maximum likelihood estimation; Little & Rubin, 2002; Schafer, 1997), minimize potential biases of traditional approaches (among them, pairwise deletion, listwise deletion, and mean substitution). Accessible articles and monographs have appeared (e.g., Allison, 2001; Schafer & Graham, 2002). Accessible software for multiple imputation (e.g., SAS PROC MI and MIANALYZE as well as the freeware NORM program; Schafer, 2006) and for full information maximum likelihood estimation has appeared as well.[6]

In measurement, we anticipate that IRT will enjoy increased presence in the psychology curriculum, as applications expand beyond ability testing to the measurement of personality traits and psychopathology. Computer adaptive testing with IRT offers the promise of briefer individ-

[5] Nearly all of the submissions to the APA quantitative methodology journal *Psychological Methods* from 2001 through 2006 were in the areas of statistics and measurement. Very few manuscripts addressing research design were submitted.

[6] SPSS includes its missing data routines in an extra cost module. Limitations of this module have been documented (von Hippel, 2004). Full information maximum likelihood estimation has been incorporated into several SEM packages, for example, AMOS, EQS, Mplus, and Mx (Neale, Boker, Xie, & Maes, 2004).

ualized tests that yield more accurate measurement of important individual differences. IRT-based techniques for assessing differential item functioning across groups can potentially help solve thorny problems of equivalent measurement in groups defined by gender, age, language, and culture. A highly accessible full introductory treatment of IRT (Embretson & Reise, 2000) has appeared. Difficult-to-use DOS-based IRT software has posed an impediment to teaching IRT. New, user friendly Windows-based IRT computer programs are currently in development and are expected to be available shortly.

Research design is noteworthy in its absence from our anticipated additions to the quantitative methodology curriculum. Only a few scattered individual scholars within psychology have research design as a main interest, and they are not concentrated at any university. This contrasts with the strong research design and methods group in psychology at Northwestern University in the 1970s.[7] The current landmark book on research design by Shadish, Cook, and Campbell (2002) was published fully 23 years following its previous incarnation (T. D. Cook & Campbell, 1979); the current version builds on the classic tradition developed by Donald Campbell. In general, psychologists are unaware of important new developments occurring in other disciplines, thus creating a vacuum of interest and demand. Indeed, many psychologists have historically conflated research design with ANOVA applied to laboratory experiments. The current lack of attention in psychology to advances in research design occurs despite the increased importance of longitudinal and field/community research in many substantive areas and despite the employment of many PhDs in applied research settings. We note that new approaches in research design typically require only existing software from commonly used statistical packages. For example, logistic regression is the basis of propensity score models used to equate participants when randomization is not possible (Rosenbaum & Rubin, 1983). SEM programs can be used for the analysis of data from experiments and observational studies with planned missingness designs (Graham, Taylor, Olchowski, & Cumsille, 2006).

Conceptualizing the Quantitative Faculty Role

Quantitative faculty role complexity. Departmental roles for quantitative faculty are complex. This role complexity poses unique challenges for the success of junior quantitative faculty. Roles for quantitative faculty can include responsibility for the yearly teaching of the introductory graduate statistics sequence to the whole psychology PhD student body, implementing a more advanced quantitative curriculum, consulting with faculty and students outside the classroom on the use of new methodologies and statistical software, serving as the methodologist on an extensive array of master's and dissertation committees in multiple substantive areas, joining research teams as methodologist, serving as methodologist on research grants, and collaborating in writing with substantive colleagues when these colleagues' manuscripts involve novel forms of complex methodology. These roles are in addition to the responsibility for the quantitative faculty member to have his or her own active research program and, in departments with quantitative PhD concentrations, to mentor quantitative PhD students. In many departments without a quantitative PhD program, the quantitative faculty member is a solo member. For solo junior quantitative faculty, there is no mentor to guide the junior faculty member through this complex array of demands and expectations. The solo quantitative faculty member also has no quantitative doctoral students who can support his or her quantitative activities and participate in his or her quantitative research program.

Supporting the quantitative faculty role. Psychology departments are challenged in two ways. First, departments are challenged to develop a model of a reasonable set of role expectations for the faculty member(s) responsible for quantitative training. Second, departments are challenged to develop mechanisms for supporting the success of young quantitative faculty members. Some departments wisely have taken steps in this direction by providing course relief for service on many student committees and for consulting or by limiting consulting hours. Some departments have turned private consulting with individuals into consulting seminars in which course credit is given to the faculty member who publicly discusses statistical and methodological issues raised by faculty and graduate students. This mechanism both limits the number of consulting hours and serves a pedagogical purpose for colleagues and graduate students. Such mechanisms offer the promise of helping overcome the service burden on junior quantitative faculty and enhancing their possibility of achieving promotion and tenure. We have observed very high mobility among junior quantitative faculty, who change jobs pretenure to find a work environment more conducive to their successful career progress. Junior quantitative faculty who endure their pretenure period under excess service demands uniquely imposed on quantitative faculty run the risk of being denied tenure because of their failure to establish an independent research program.

Continued Human Resources for Quantitative Training—A Multilevel Challenge

Many of the faculty teaching statistics, measurement, and research design are at or nearing retirement age. As these faculty retire, there are numerous potential ramifications—the quality of graduate training in quantitative methodology for substantive psychologists, the production of new PhDs in quantitative psychology, the methodological support for substantive research projects and grant proposals, and even the quality of the methodological peer review of manuscripts submitted for publication are all at risk of decline (Clay, 2005).

The quantitative methods training of future generations of graduate students presents a difficult challenge for

[7] The Northwestern University group included Donald Campbell, as well as Robert Boruch, Thomas Cook, Albert Erlebacher, Lee Sechrest, Benton Underwood, and numerous graduate students and postdoctoral fellows with interests in methodology.

psychology. Employment opportunities for quantitative psychologists within and beyond the academic setting will exceed the remarkably small supply of new quantitative PhDs. Elite psychology departments report having conducted multiyear searches in an attempt to replace a retiring or departing faculty member in quantitative methods.

The broad range of quantitative faculty. As our survey showed, there is a broad continuum of faculty members involved in quantitative training, both in terms of their own quantitative backgrounds and their research and scholarship. At one end is a small group of quantitative methodologists or developers (B. Muthén, 1989) whose PhD training is in quantitative methods per se. This group creates and evaluates new methodologies and publishes work in technical quantitative outlets. At the other end of the continuum is a much larger group of "bridgers, who are well trained quantitatively but do not publish original methodological or quantitative work. These bridgers are typically individuals in a substantive area who have pursued quantitative training beyond that required for their substantive degrees" (Aiken et al., 1990, p. 733). In between are faculty who combine quantitative and substantive scholarly careers to one degree or another.

Two classes of workforce needs. We parse the issue of the future of quantitative training into two broad issues. The first is the training of the wide range of new PhDs in all areas of substantive psychology. Second is the training of a new generation of quantitative faculty, the developers who will advance the field of quantitative psychology.

We first address the issue of quantitative training for substantive psychologists. Faculty who combine substantive psychology with quantitative training are often considered highly desirable hires. We term these individuals *twofers* (two for the price of one), faculty who can contribute to quantitative training of doctoral students across psychology and at the same time strengthen a substantive area. This model poses extraordinary challenges to the young faculty member who must meet the quantitative needs of the department while simultaneously building his or her own research program and participating in training and curricular innovation in the substantive area.

We can immediately identify some giants, exemplars who have been both powerful methodologists and great substantive psychologists, both in the past and at present. Among those who have gone before are Lee J. Cronbach, Raymond B. Cattell, and Julian C. Stanley. However, we cannot expect this to be the norm, and even some of the giants made their methodological and substantive contributions during different parts of their careers. Can average or even strong twofers meet the demands of their substantive career foci and also stay abreast of new quantitative methodology over time and implement new quantitative methodology courses? This question is particularly significant for those twofers who serve as the solo quantitative psychologist in a department. Over time, do twofers leave quantitative teaching, drawn much more to their substantive careers and away from quantitative service teaching? Alternatively, do twofers become the faculty members who teach the introductory statistics course over and over throughout their careers? These questions bear examination as we consider the maintenance of the quantitative professoriate in psychology.

Next, we address the issue of training quantitative psychologists. The training of doctoral students in quantitative psychology is a matter of serious concern. Mentoring those who will produce new quantitative methodologies for psychology must be accomplished by quantitative psychologists who themselves are actively involved in both research and scholarly publication in quantitative methodology. Quantitative psychology is no different from any other area of psychology in this regard. One concern is whether the emphasis on the twofer model of faculty who teach quantitative methods but whose scholarship is substantive undermines the training of future quantitative psychologists.

Short-Term and Long-Term Faculty Solutions

The short-term perspective. In the short-term, one resource for enhanced quantitative training for all doctoral students must be existing faculty. Departments should examine whether they are making the best use of their quantitative faculty. We note that 81% of faculty who were teaching graduate-level quantitative courses at the time of our survey were also teaching substantive courses. Some of them could be released from substantive teaching obligations to teach new quantitative curriculum. In some departments, quantitative faculty are limited to teaching the large introductory graduate statistics sequence course each year, precluding them from the opportunity to implement new curriculum. In other departments, quantitative faculty are often found teaching undergraduate quantitative methods courses, which is good from the perspective of enhancing undergraduate training in quantitative methods but is to the detriment of PhD training. If new courses are to be implemented, retraining may well be required, and faculty will need time to retrain and to develop the quantitative curriculum.

A second resource is quantitative training in other departments on campus. Under the best of circumstances, faculty in psychology become fully aware of training resources across campus, interact with other quantitative faculty on campus, and informally audit courses in other departments to judge their appropriateness for graduate students in psychology. One concern with quantitative courses taught outside psychology is whether these courses address research issues in psychology (e.g., the much smaller samples in psychology than in related social sciences and education). A second concern is whether the quantitative material is contextualized in a manner useful to psychology students (e.g., the application of current measurement techniques to personality and psychopathology indices rather than to ability assessment). Lovett and Greenhouse (2000) showed the importance of presenting statistical concepts in the context of the substantive area in which they will be applied. If a department makes use of outside-of-department courses, it may be useful for faculty from psychology to supplement the examples from these courses with examples from psychology. Once a decision is made to make other departments responsible for the statistics and methods education of graduate students in psychology, it is often a difficult decision to undo—the higher administration may have committed resources to the other department to support this training effort, and psychology may have in turn lost the potential to gain parallel resources.

The long-term perspective. In the long term, psychologists have to rebuild the infrastructure for quantitative methodology training within psychology. We must increase awareness within our own discipline of the nature of quantitative methodology as a research subdiscipline of psychology. Mathematically talented students at all levels must be encouraged to consider quantitative psychology as a field. Departments of psychology should work to strengthen their own quantitative infrastructures and to ensure that existing quantitative training programs continue and remain strong. Support should be given to the building of new quantitative programs. To this end, in 2006 the APA funded the Task Force to Increase the Quantitative Pipeline. The report of this task force focuses on how to increase awareness of quantitative psychology as a discipline and on how to recruit talented individuals to the field (APA Task Force to Increase the Quantitative Pipeline, 2007).

What is the view of quantitative research in psychology? Even a cursory review of current job advertisements for faculty to teach advanced quantitative methods reveals that departments often seek to hire faculty who have a substantive rather than a quantitative research program. Apparently, quantitative methodology as a subdiscipline of psychology is not well understood as a distinct area of research and scholarship for a faculty member in psychology. Perhaps this is because the role of the quantitative faculty member is viewed as a service role, rather than as a legitimate faculty role for a scholar in psychology. Yet, methodological research has a broad and lasting impact.[8] The future of quantitative methods in psychology requires that the broad discipline of psychology come to understand quantitative methodology as a field of scholarly endeavor and to support its continuation within psychology.

Other Mechanisms for Training of Graduate Students and Retraining of Quantitative Faculty

Access to quantitative training beyond the bounds of home departments for both doctoral students and for faculty is a critical issue. A number of universities are experimenting with various mechanisms for making advanced quantitative methods training available to their graduate students. Some universities have utilized distance learning methods, through live interactive broadcasts to multiple campuses or through live or recorded streaming video. A second mechanism is the use of workshops in measurement and statistics. These workshops vary in length from 1 to 2 hours to a week or longer, vary in level from introductory overviews to advanced training institutes, and vary in focus from learning to use a specific computer program to a broad introduction to a quantitative technique. A rarely used potential training mechanism is for students to take advanced courses at another university for part of their graduate training. A small number of postdoctoral mechanisms exist for strengthening methodological training; these are primarily limited to U.S. citizens. A final training mechanism is the use of online courses. These courses are often implemented with faculty in statistics, business, engineering, and medicine. For example, the American Statistical Association also offers online continuing education for faculty in addition to more traditional workshops (American Statistical Association, 2007).

Such methods certainly make presentations of statistical material available to a wider audience at low cost. To our knowledge, psychology has not undertaken a systematic evaluation of these alternative teaching approaches to enhancing competency in quantitative methodology. Such an evaluation is warranted. Areas of challenge include matching the instructional material to the participant background and providing assistance at the home institution to support the application of workshop content. We believe that such mechanisms as distance learning, workshops, and online courses can supplement but not supplant adequate instruction within graduate departments of psychology.

What Is an Optimal Level of Quantitative Training for PhDs in Psychology?

The characterization of the optimal level of quantitative training poses an interesting and complex challenge. Quantitative curriculum is the last core curriculum in psychology. A question to be considered then is whether there is a fundamental body of knowledge in quantitative methods that everyone should master. Back in the 1960s and early 1970s, the body of quantitative knowledge included factorial design and ANOVA for experimental and correlational psychologists (Cronbach, 1957). Correlational psychologists additionally learned factor analysis, classical test theory, and test construction.

Now, very large segments of psychology, including those formerly distinguished as experimental versus correlational, need to know a great deal about multivariate data structures and about latent variables. Every psychologist who will make use of an existing scale or will invent a scale should have fundamental measurement training, including IRT. Training in research design needs to incorporate designs for field settings, observational studies, longitudinal studies, and studies in which random assignment is not possible. Beyond this basic core, there are certainly areas

[8] During the period when both substantive and methodological articles were published in *Psychological Bulletin*, 7 of the top 10 *Psychological Bulletin* citation classics were methodological in nature (Sternberg, 1992). The impact factor of the APA journal *Psychological Methods* averaged 4.17 from 2003 through 2005, equivalent to or exceeding that of the major substantive outlets in the APA family of journals: *Journal of Clinical and Consulting Psychology* (4.02), *Journal of Personality and Social Psychology* (3.90), *Developmental Psychology* (3.12), *Behavioral Neuroscience* (2.93), *Journal of Experimental Psychology: Human Perception and Performance* (2.77), *Journal of Experimental Psychology: Learning, Memory, and Cognition* (2.45; ISI Web of Knowledge, 2007).

within psychology in which further specialized quantitative training is required.

Conclusion

Our survey highlighted areas of gain in graduate training in measurement and statistics in PhD programs in psychology over a dozen years. Yet, it also identified important areas of training with little or no improvement. Quantitative training lagged behind in transmitting important innovations that would potentially serve to enhance the future of the science of psychology. We continue to have great concern about training in measurement. We are also profoundly troubled about the area of research design, which appeared stagnant with some decline in curriculum and competency. With the imminent retirement of many senior quantitative faculty during the next decade, psychology programs are challenged to recruit PhD students with methodological interests into psychology and to create curricular structures in which they can train intensively in quantitative methods, typically while also pursuing their substantive interests. Beyond this, steps must also be taken within academic settings to nurture new PhDs hired to bear the responsibility of carrying out and enhancing the quantitative curriculum so that they can successfully achieve promotion and tenure. These steps must be taken to maintain and improve the quantitative training of PhDs in psychology in the service of advancing our science.

REFERENCES

Aiken, L. S., West, S. G., Sechrest, L. B., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology: An international survey. *American Psychologist, 45,* 721–734.

Allison, P. D. (2001). *Missing data.* Thousand Oaks, CA: Sage.

American Psychological Association. (1997). *Graduate study in psychology, 1997.* Washington, DC: American Psychological Association.

American Psychological Association Task Force to Increase the Quantitative Pipeline. (2007). *Report of the Task Force to Increase the Quantitative Pipeline: Draft final report.* Washington, DC: American Psychological Association.

American Statistical Association. (2007). *Center for Statistics Education: Graduate, professional, and continuing education.* Retrieved May 15, 2007, from http://www.amstat.org/education/index.cfm?fuseaction=graduate

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91,* 444–455.

Barnard, J., Du, J. T., Hill, J. L., & Rubin, D. B. (1998). A broader template for analyzing broken randomized experiments. *Sociological Methods and Research, 27,* 285–317.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31,* 419–456.

Bentler, P. M. (1985). *Theory and implementation of EQS, A structural equations program.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1995). *EQS program manual.* Los Angeles: Multivariate Software.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

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

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., Seltzer, M., & Congdon, R. (1988). *An introduction to HLM: Computer program and user's guide* (2nd ed.). Chicago: University of Chicago Department of Education.

Clay, R. (2005). Too few in quantitative psychology. *Monitor on Psychology, 36*(8), 26–28.

Cook, R. D., & Weisberg, S. (2004). ARC [Computer software]. Retrieved from http://www.stat.umn.edu/arc/software.html

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

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

Embretson, S. E., & Reise, S. R. (2000). *Item response theory for psychologists.* Mahwah, NJ: Erlbaum.

Goldberger, M. L., Maher, B. A., & Flatteau, P. E. (Eds.). (1995). *Research-doctorate programs in the US: Continuity and change.* Washington, DC: National Academy Press.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11,* 323–343.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances.* Baltimore, MD: Johns Hopkins University Press.

Herszenhorn, D. M. (2006, May 5). As test-taking grows, test-makers grow rarer. *New York Times,* pp. A1, A19.

Hox, J. (2002). *Multilevel analysis: Techniques and applications.* Mahwah, NJ: Erlbaum.

Insightful Corporation. (2007). S-PLUS [Computer software]. Seattle, WA: Author.

ISI Web of Knowledge. (2007). *Journal citation reports.* Retrieved May 13, 2007, from http://portal.isiknowledge.com

Kenny, D. A. (1979). *Correlation and causality.* New York: Wiley.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling.* Thousand Oaks, CA: Sage.

Lambert, N. M. (1991). The crisis in measurement literacy in psychology and education. *Educational Psychologist, 26,* 23–25.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Loehlin, J. C. (1987). *Latent variable models: An introduction to factor, path, and structural analysis.* Hillsdale, NJ: Erlbaum.

Lovett, M. C., & Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *American Statistician, 54,* 196–206.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* New York: Erlbaum/Taylor Francis.

Meier, S. T. (1993). Revitalizing the measurement curriculum: Four approaches for emphasis in graduate education. *American Psychologist, 48,* 886–891.

Merenda, P. F. (1996). Note on the continuing decline of doctoral training in measurement. *Psychological Reports, 78,* 321–322.

Merenda, P. F. (2003). Measurements in the future: Beyond the 20th century. *Psychological Reports, 92,* 209–217.

Merenda, P. F. (2006, August). *Psychometrics and psychometricians in the 20th and 21st centuries: How it was in the 20th century and how it is now.* Invited paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Millsap, R. E., & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. In R. C. MacCallum & R. Cudeck (Eds.), *Factor analysis at 100* (pp. 131–152). Mahwah, NJ: Erlbaum.

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.

Muthén, L. K., & Muthén, B. O. (1998–2007). *Mplus user's guide* (4th ed.). Los Angeles: Authors.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2004). Mx: Statistical modeling [Computer software]. Richmond, VA: Virginia Commonwealth University, Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry.

Norusis, M. J. (2002). *SPSS 11.0 guide to data analysis.* Saddle River, NJ: Prentice Hall.

Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit hierarchical linear and growth trajectory models. *Educational and Psychological Measurement, 65,* 811–835.

Rasbash, J., Steele, F., Browne, N. J., & Prosser, B. (2005). *A user's guide to MLwiN. Version 2.0.* Bristol, England: University of Bristol, Centre for Multilevel Modeling. Retrieved from http://www.cmm.bristol.ac.uk/MLwiN/index.shtml

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rosenbaum, P. R. (2002). *Observational studies* (2nd ed.). New York: Springer-Verlag.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41–55.

R Project for Statistical Computing. (2007). R [Computer software]. Retrieved from http://www.rproject.org

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association, 100,* 322–331.

SAS Institute. (1992). The MIXED procedure. In *SAS Technical Report P-229, SAS/STAT software: Changes and enhancements, Release 6.07* (pp. 287–366). Cary, NC: Author.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* New York: CRC Press.

Schafer, J. L. (2006). *Norm* [Computer software]. Retrieved October 23, 2007, from http://www.stat.psu.edu/~jls/misoftwa.html

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shadish, W. R., Luellen, J. K., & Clark, M. H. (2006). Propensity scores and quasi-experiments: A testimony to the practical side of Lee Sechrest. In R. R. Bootzin & P. E. McKnight (Eds.), *Strengthening research methodology: Psychological measurement and evaluation* (pp. 143–157). Washington, DC: American Psychological Association.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24,* 322–354.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

SPSS. (2001). SPSS 11.0 advanced models [Computer software]. Saddle River, NJ: Prentice Hall.

SPSS. (2007). *Amos 16.0 user's guide.* Chicago: Author.

Sternberg, R. J. (1992). *Psychological Bulletin*'s top 10 "hit parade." *Psychological Bulletin, 112,* 387–388.

Vellman, P. F. (2006). *DataDesk user's manual.* Ithaca, New York: Data Description.

von Hippel, P. T. (2004). Biases in SPSS 12.0 missing value analysis. *American Statistician, 58,* 160–164.

West, S. G., & Sagarin, B. J. (2000). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy* (Vol. 2, pp. 117–154). Thousand Oaks, CA: Sage.

West, S. G., & Thoemmes, F. (in press). Equating groups. In P. Alasuutari, J. Brannen, & L. Bickman (Eds.), *Handbook of social research methods.* London: Sage.

Widaman, K. F., & Reise, S. P. (1997). Exploring the invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.

Zimowksi, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2006). BILOG-MG 3 [Computer software]. Chicago: Scientific Software International.

(*Appendix follows*)

Appendix
Methodological Procedures of the Survey

Procedure and Respondents

In April 1998, we mailed questionnaires to the chairs of all 234 programs identified in *Graduate Study in Psychology, 1997* (American Psychological Association, 1997) as conferring the PhD in at least one area of psychology. Chairpersons were asked to identify the faculty member most knowledgeable about the methodology curriculum to complete the questionnaire. We sent reminder letters four weeks later. In September 1998, we again sent questionnaires to chairpersons of nonresponding departments, with reminder letters sent one month later. We simultaneously made multiple personal contacts by letter, e-mail, and telephone with faculty at these institutions who might respond. Data collection terminated in spring 1999, with 201 completed questionnaires (an 86% overall response rate).

Elite Institutions

We used the 1995 National Research Council quality rankings (Goldberger et al., 1995) to identify elite institutions, which we defined as the top 25 ranked departments. In all, 22 of these 25 departments (88%) responded to our survey, and 44 (88%) of the top 50 departments responded, mirroring the overall response rate.

Substantive Areas Represented Among Responding Departments

Of all programs, 75% (73% elite) had a clinical area; 61% (95% elite) had a cognitive area; 60% (95% elite) had a social area; 55% (92% elite) had a biopsychology or neuropsychology area; 53% (77% elite) had a developmental area; 30% (36% elite) had an industrial–organizational, engineering, or human factors area; 24% (32% elite) had a personality area; and 9% (28% elite) had a counseling area.

Questionnaire

The questionnaire addressed the following areas: (a) program demographics, including total number of PhD students and faculty, the number of faculty teaching any aspect of methodology at the PhD level, the use of adjunct or part-time faculty for methodology training, the use of out-of-department methodology curriculum in PhD training, and the existence of a quantitative PhD concentration within the department; (b) the full departmental curriculum in statistics, measurement, and research design, measured in terms of the frequency with which each of a large variety of topics was taught in the psychology curriculum, the duration of coverage of each topic, and the availability of appropriate-level coverage of the topic for psychology PhD students elsewhere on campus; (c) topics added to the methodology curriculum over the past decade and topics still lacking in the curriculum; (d) doctoral-level introductory statistics course (or course sequence), including whether such a course existed, whether it was taught in the department, the duration of the course, whether the course was universally required of all students, and a detailed account of course content; (e) availability of a measurement course for all PhD students; (f) availability of a general research design course for all PhD students, plus research methods courses in individual substantive areas; (g) the number of required statistics and measurement courses by substantive area; (h) training in computing and computing resources; (i) judged competency of new PhDs to apply a variety of methodologies in statistics, measurement, and research design to their own work; (j) backgrounds and interests (quantitative vs. substantive) of faculty teaching PhD-level quantitative methodology courses and turnover in those faculty teaching methodology; (k) creation and dissolution of quantitative PhD programs and the numbers of PhD students being trained in quantitative methods.