Research in Higher Education, Vol. 45, No. 4, June 2004 (© 2004)

A REVIEW OF METHODOLOGICAL CHARACTERISTICS OF RESEARCH PUBLISHED IN KEY JOURNALS IN HIGHER EDUCATION: Implications for Graduate Research Training

Susan R. Hutchinson*,† and Cheryl D. Lovell**

Professional journals serve an important function within most disciplines as they offer a mechanism for professional communication. In the field of higher education, research on methodological characteristics of the published literature has been sparse. This study used content analysis to identify the types of research designs and analytical approaches utilized most often in 3 leading higher education journals during 5 recent years. Results indicated that across all 3 journals, most studies used quantitative analyses based on either primary or secondary survey data, though there were some differences among the journals. The bulk of quantitative studies were based on procedures normally taught in intermediate and advanced statistical courses. The results suggest that graduate-level research training may need modifications if higher education professionals are to contribute to or fully appreciate the published higher education literature.

KEY WORDS: higher education; research training; research methodology; journal quality; journal characteristics; journal content.

INTRODUCTION

Professional journals serve an important function within most disciplines. They offer a mechanism by which professionals communicate ideas, stimulate discussion (as well as controversy), and share information, often in the form of research findings. This latter role is well known in many fields, such as medicine (Hummel, 1984) and nursing (Love, 1996), which rely on journals as a primary source of continuing professional education. In the field of higher education, Carpenter (2001) argues that keeping current with the published literature

*University of Northern Colorado, Greeley, CO.

**University of Denver, Denver, CO.

†Address correspondence to: Susan Hutchinson, Department of Applied Statistics and Research Methods, Box 124, University of Northern Colorado, Greeley, CO 80639. E-mail: susan.hutchinson@unco.edu

0361-0365/04/0600-00383/0

comprises part of student affairs' ethics: "Any student affairs worker who does not read the literature, does not stay current professionally through continuing education, . . . is not acting ethically" (p. 306). However, as Kuh, Bean, Bradley, and Coomes (1986) and Goodwin and Goodwin (1985) have noted, the utility of journals as a source of current information is often limited by the research training of the reader. To the extent that journals report research methods and procedures not typically covered in graduate-level training, the information presented is unlikely to be appreciated or fully understood. This may be why, in part, higher education practitioners and researchers have reported that journals in higher education tend to be difficult to read and of minimal usefulness (Kezar, 2000). Malaney (2002), too, has observed the deficiency in methodological training among higher education and student affairs professionals, which may limit not only their ability to produce research but also their ability to comprehend the existing, published research.

Numerous reviews of professional journals have been conducted for the purpose of classifying studies in terms of their designs and methodological characteristics. For example, Goodwin and Goodwin (1985) coded statistical techniques appearing in the *American Educational Research Journal* (*AERJ*) between 1979 and 1983. They concluded that graduate students with training through the intermediate statistics level and with knowledge of basic psychometric principles would be able to comfortably read approximately 80% of the studies in *AERJ*. In a followup to the Goodwin and Goodwin study, Elmore and Woehlke (1998) determined that the relative use of various types of quantitative methods in *AERJ* has remained fairly constant since 1983, although use of qualitative research methods has increased over the past 10 years. Similar findings were reported by Baumberger and Bangert (1996) in an analysis of research designs and statistical techniques reported in the *Journal of Learning Disabilities* between 1989 and 1993, although they did not report on the use of qualitative methods. More recently, Kieffer, Reese, and Thompson (2001) examined methodological characteristics of the *AERJ* and the *Journal of Counseling Psychology* (*JCP*) over a 10-year period spanning 1988–1997. Their purpose was to examine analytic features of quantitative studies in relation to recommendations offered by the American Psychological Association's Task Force on Statistical Inference (Wilkinson and APA Task Force, 1999 as cited in Kieffer et al., 2001) regarding various aspects of data analysis and reporting of such analyses. They concluded that overall, researchers in both journals used univariate statistics more often than multivariate and that usage of particular statistical procedures was quite stable across the 10-year period.

In the field of higher education, research on methodological characteristics of the published literature has been sparse. Kuh, Bean, Bradley, and Coomes (1986) and Kuh, Bean, Bradley, Coomes, and Hunter (1986) conducted two

studies in which they examined journals specifically targeted toward student affairs professionals as well as several other journals that included articles on college students between 1969 and 1983. They found the research in the student affairs journals to reflect primarily descriptive studies, with less than a third reporting any type of multivariate technique (including multiple regression; Kuh, Bean, Bradley, Coomes, and Hunter, 1986). Moreover, when comparing the student affairs journals to the other journals, they found the other journals used more multivariate as well as data reduction procedures. Although the studies by Kuh and his associates were fairly comprehensive, they were confined to articles specifically concerning college students.

Information about the methodological characteristics of the higher education literature in general, and in particular, about the research published in several leading higher education journals, *The Journal of Higher Education* (*JHE*), *Research in Higher Education* (*ResHE*), and *Review of Higher Education* (*RevHE*) is generally lacking. Although Kuh et al. included these journals in their two studies (Kuh, Bean, Bradley, and Coomes, 1986; Kuh, Bean, Bradley, Coomes, and Hunter, 1986), their use of aggregated findings precludes scrutiny of the individual journals. Volkwein, Carbone, and Volkwein (1988) provided limited examination of trends in methodologies appearing in *ResHE* across 15 years from 1973 to 1987. They concluded that methodological characteristics were relatively stable across the 15-year publication period with most (over 50%) studies reporting some type of multivariate statistical procedure and few (5%) representing historical or case studies. However, the focus of their study appeared to be more about the article topics and author characteristics than about the methods. As a result, their study provided little detail concerning the nature of the methods used in the coded studies such as data sources, designs, and types of statistical procedures applied. Milam (1991) and Silverman (1987) both conducted reviews that used these three journals, but Milam coded the journals on the basis of paradigmatic characteristics and not methodology *per se*, while Silverman only categorized the articles by topic. Consequently, little is known about the designs or methodological approaches presented in these three higher education journals. Yet, given recent attention to scholarship and graduate training in higher education and student affairs (see e.g., special issues of the *NASPA Journal* [Roper, 2002] and *Journal of College Student Development* [Blimling, 2001]), examination of the methodological features of *JHE, ResHE*, and *RevHE* should be of interest to higher education professionals in general and to those responsible for higher education graduate training programs in particular.

The present study used content analysis to identify the types of research designs and analytical approaches utilized most often during 5 recent years of *JHE, ResHE*, and *RevHE*. In addition, we were interested in comparing the three journals to determine if studies reported in the three journals differed in their

research approaches and levels of methodological sophistication. The findings are discussed in terms of implications for graduate-level research training in the field of higher education.

METHODS

Data Source

The three journals reviewed for this study, the *Journal of Higher Education, Review of Higher Education*, and *Research in Higher Education*, were chosen because, as others have noted (Milam, 1991; Silverman, 1987), they are among a core of higher education journals that enjoy relatively high prestige. Each journal has a board of external reviewers and utilizes a blind review process. A low acceptance rate (or high rejection rate) is often considered an indicator of quality and indeed each of the journals has a relatively low acceptance rate with the *JHE* accepting only 6%–10% (Cabell and English, 2000b), *ResHE* accepting 11%–20% (Cabell and English, 2002a), and the *RevHE* accepting 21%–30% (Cabell and English, 2002b) of the manuscripts reviewed. Both *ResHE* and the *RevHE* are sponsored by professional associations and tend to be focused toward faculty and students involved in the study of higher education, academics in general, and college and university administrators engaged in a variety of professional responsibilities. Of the three, only the *JHE* is published by a university press, though it equally addresses a broad array of issues within the higher education context. Both the *JHE* and *ResHE* are published six times a year while the *RevHE* is published quarterly. These journals are primary reading for individuals interested in the study of higher education. Consequently, we believed they represent the state of the art with respect to research being conducted in the field.

We restricted our search to 5 of the most recent years (1996–2000), in part because our study was cross-sectional in nature and because prior research has suggested that methodological approaches tend to be stable within a 5-year span (Goodwin and Goodwin, 1985). Therefore, we felt confident that a 5-year sample would provide an adequate representation of current research in higher education. After data collection, χ^2 tests of independence provided empirical support that the three journals did not differ significantly ($p < .01$) between 1996 and 2000 on several key variables, including type of article, design type, and level of statistical analysis reported. To examine the possibility that the consistency of article characteristics across the 5-year coding period might be an artifact of stable editorial board memberships, we computed the proportion of board members who remained on their respective editorial boards across the entire coding period. While the editors for all three journals were the same throughout the coding period, board membership changed considerably, with none of the

JHE members remaining on their editorial board from 1996 to 2000 and fewer than 12% of the *RevHE* board remaining on the board during that period. The editorial board for *ResHE* was the most stable, with 53% of the 1996 board still in term at the end of 2000. We coded all articles except book reviews and editorials, with one notable exception. Because the September/October 1999 issue of *JHE* was a special anniversary issue presenting previously published articles from the 1930s, we made a decision to omit the entire issue as it seemed somewhat anomalous and inconsistent with our purpose of documenting current research methods. Four hundred and two ($N = 402$) articles were coded for the study, with 32.1% ($n = 129$) from *JHE*, 25.9% ($n = 104$) from *RevHE*, and 42.0% ($n = 169$) from *ResHE*.

Instrumentation

Development of the Coding Form

We developed a form for the purpose of conducting content analysis of the articles in our sample because no existing coding sheet included all of the information we were seeking. The form was partitioned into the following sections: general identifying characteristics of the article (e.g., journal name, year, title), research design, data collection methods/source, types of statistical analyses used, and quality of article in reporting measurement and sampling information. We included this latter section as a way to determine the sophistication of the researchers with respect to measurement concepts, as measurement training is often given minimal attention in graduate education (Aiken, West, Sechrest, and Reno, 1990; Lovell, Hutchinson, and Fairweather, 1999). An initial draft of the form was developed based partly on examination of other studies that had employed similar content analyses of research articles. For example, our list of statistical techniques was a modification of the list used by Goodwin and Goodwin (1985). As Goodwin and Goodwin had done, we also coded statistical techniques as basic, intermediate, or advanced, although our categorization differed somewhat. The purpose of coding these techniques in our study was to examine the level of the published literature relative to the research training most higher education doctoral students receive. We defined *basic* statistics as those typically covered in a first semester, introductory-level statistics course. *Intermediate*-level statistics were those generally taught in a second or third semester course, for example, in an ANOVA or multiple regression class. *Advanced* statistical procedures were defined as those that would normally require a fourth course or beyond in statistics or research methods. For the most part, procedures characterized as advanced require training in multivariate statistics or in specialized methodological areas such as structural equation modeling or measurement theory. Our design categories were taken largely from the scheme presented by

Isaac and Michael (1984). To assess clarity and completeness of the coding form, we independently scanned one year's worth of each of the three journals to look for salient methodological features that might have been inadvertently omitted from our form. As a result minor modifications were made.

To assess reliability of the coding scheme, we both coded all articles in one randomly selected issue of *RevHE*. Overall interrater agreement based on simple percentage agreement was 92%. Although there was 100% agreement across all 22 statistical techniques, there was some disagreement on the higher inference items, particularly those items used to code "quality." As a result, one item on sampling was deleted from the coding scheme, and a more specific set of operational definitions was created for those items that had led to greatest disagreement. Interrater agreement improved to 96% overall when the item regarding sampling was omitted. The final version of the coding form is in the appendix.

Operational Definitions

The criterion for coding the statistical procedures listed on the coding form was a relatively low-inference task based on the simple dichotomy: either the study reported using the procedure or the study did not. Although authors did not always explicitly identify the particular statistical test used, in most cases the specific procedure could be fairly easily deduced, for example, through context or examination of degrees of freedom. As mentioned previously, interrater agreement on coding of statistical procedures was quite high (100%), thus supporting the generally unambiguous nature of coding use of particular statistics. Similarly, coding of data collection source was relatively straightforward and easily determined through the authors' description of their methods: the author conducted interviews or did not, the author examined documents/artifacts or did not, and so on. Primary survey was coded if a survey was developed and data collected specifically for the study being reported, whereas secondary survey was coded if the authors used data collected for a different purpose or from a different study.

For other nonstatistical elements of each article, such as type of article, design used, and quality of article, we created specific operational definitions to guide the coding process. These definitions, presented below, were developed in part based on standard research methods texts (Cook and Campbell, 1979; Creswell, 2002; Gall, Gall, and Borg, 2003; Krathwohl, 1998) as well as from the authors' research experience and expertise. Both authors have considerable training and experience in research. The first author is a quantitative research methodologist, and the second author is a postsecondary policy analyst.

Articles were coded into one of six possible types including expository, expository supported by data, opinion/position paper, qualitative study, quantitative empirical, and historical review. For the purposes of data analysis, however,

these six categories were collapsed into four: expository/opinion, expository supported by data, original research (either qualitative or quantitative), and historical review. Expository articles were those addressing an issue primarily through review of literature. Examples might include proposing a conceptual model or offering strategies for improving retention. Opinion/position articles primarily reflected views of the authors and appeared to be of an editorial nature, though these articles were not identified as editorials, *per se*. Expository supported by data was a category we created for articles that did not seem to fall clearly into either the expository or empirical categories. These were articles that were primarily expository in nature but with some data (e.g., tables or graphs from secondary sources) used to support the authors' argument. What distinguished these from either a qualitative or quantitative study was that the article provided no evidence of data collection, for example, did not include description of data collection or data analysis procedures. Supportive data tended to be cited from other sources. Historical review referred to articles describing some past phenomenon but that had no stated design or methodology. In essence, historical reviews were expository works but with a historical perspective. Historical reviews differed from qualitative studies based on historical design in that the latter presented clearly delineated data sources and methods; the former did not appear to involve original data collection or analysis. Both qualitative and quantitative studies were distinguished from other article types by their inclusion of evidence that the article presented original research. Such articles included explicit descriptions of participants/data sources, data collection methods, and data analysis procedures.

Regarding the design categories, for "traditional" quantitative designs such as true experiment, quasi-experiment, correlational/passive observation, and ex post facto/causal comparative, we relied on standard definitions from Cook and Campbell (1979), Gall et al. (2003), and Krathwohl (1998). We operationally defined descriptive quantitative studies as those in which no inferential statistical tests were used. Although most of the quantitative studies reported some type of descriptive information—means, percentages, and so on—they were not classified as descriptive designs unless description was the sole or primary purpose of the study. Among studies using qualitative methods, we coded studies as qualitative descriptive unless their authors specified a particular type of design, for example, ethnography or narrative inquiry, or unless the specific design was clear from the methodological description. For instance, if in-depth examination of organizational culture was the purpose of a study, the study was coded as an ethnographic design even if the author did not call the study ethnographic. In most cases, authors did not refer to a particular design but merely indicated they were employing qualitative methods. Case studies (which included both qualitative and quantitative approaches) were studies of a single instance (e.g., a single academic department, single institution, etc.) in which the single in-

stance was the focus of the study. In some cases, multiple case studies were coded when several instances (e.g., multiple institutions) were specifically selected for study. Studies were coded as developmental/longitudinal when change across time was of primary interest. Longitudinal studies were fairly obvious by their use of data collected at multiple points in time and their analysis of change across those different time periods. Studies classified as historical were those in which a particular time period or past event was studied. Historical studies generally relied heavily on documents and interviews as data sources. Studies using action research were characterized by their application of research findings to a particular, practical problem. In most cases, authors of action research studies identified them as such. Meta-analyses were studies in which researchers aggregated and applied statistical procedures to a collection of prior studies on a particular topic. Test or scale validation studies were usually described as such by the authors in their statement of purpose.

The final section of the coding form was used to rate certain aspects of the quality of the articles. Items were included to determine if authors discussed statistical assumptions, provided clear operational definitions of variables, presented reliability evidence, presented validity information, and for survey studies, addressed potential nonrespondent bias. In cases where the category was not relevant we simply treated codes as missing data; for example, in a qualitative study, discussion of statistical assumptions would not apply. For each of the categories that were relevant to a particular study, we coded *no* if the author did not mention anything at all about that aspect of the study. Statistical assumptions were coded as *yes, minimal* if authors discussed or assessed tenability of one or two assumptions, and were coded as *yes, thorough* if they discussed and/or tested most or all assumptions. Regarding provision of clear operational definitions of variables, our principal guiding criterion was whether or not definitions were sufficiently detailed so that another researcher could replicate the study with respect to measurement of the variables. If some definitions were missing or if definitions were present but vague, we coded *yes, partial* and *yes* otherwise. Presentation of scale reliability was only coded as *yes* if reliability estimates were provided for all multiple-item measures based on the current sample. If reliability was provided only for some variables or was based on prior studies even if provided for all variables, then it was coded as *yes, partial or past.* Where only single-item measures of variables were reported, we coded *no scale used.* We coded presentation of validity information as *yes*, *thorough* only if multiple types of validity evidence were presented. Where only a single type of validity evidence was presented (e.g., review of item content by a panel of experts), we coded validity as *yes, minimal*. For studies using either primary or secondary survey data, we examined the article to determine if the authors addressed the potential problem of nonrespondent bias. Articles received a *yes*

even if they only mentioned possible lack of representativeness of respondents as a problem. Articles were coded as *no* if they did not mention possible nonrespondent bias at all, unless they also reported a 100% response rate.

Coding Procedures

We divided the total sample of articles between us so that each reviewer coded articles from each of the three journals across different years. Despite our fairly high interrater consistency, we did not want to impose potential bias onto our data by having a single rater rating all of a particular journal. A separate coding form was used for each journal article. Our coding procedures were fairly standard for quantitative content analysis (see e.g., Gall et al., 2003). For each article the rater read the article and examined it for information related to each coding category on the coding form (see appendix). For example, after entering the basic descriptive information for the article (i.e., journal time, year, article title, etc.) the rater determined into which broad article type the article fell. For articles representing either qualitative or quantitative studies, the rater then considered what type(s) of research design(s) was/were used in the particular article. This process continued until all relevant coding categories were completed for all of the articles in our sample.

Data Analysis

To determine the relative use of different research designs, data collection sources/methods, and statistical analyses reported across the three journals, the multiple response procedure within the SPSS statistical package (Windows version 11.0) was used to obtain frequencies and percentages. This procedure is appropriate when frequencies are based on categories that are not mutually exclusive, as was the case in this study. For example, it was quite common for a single study to incorporate multiple data sources, various types of statistical analyses, and even two or more research designs. A limitation of the multiple response procedure is that it is descriptive only and does not permit inferential hypothesis testing. When testing whether or not relative use of specific designs, methods, and data analysis procedures was comparable across the three journals, χ^2 tests of independence were conducted.

During the coding process it became apparent that a subset of authors had reported multiple studies from the same data sets (although we did not document exact numbers), often repeating verbatim their descriptions of methods and procedures. Because we were concerned about the potential violation of the independence of observations assumption when conducting statistical tests, we opted to use a somewhat conservative alpha level of .01 to reduce the risk of Type I error.

RESULTS

Of the 402 articles coded for this study, 77.6% ($n = 312$) represented original research, 12.4% ($n = 50$) were expository/opinion only, 7.7% ($n = 31$) could be characterized as primarily expository with some data support, and 2.2% ($n = 9$) comprised historical reviews. However, the relative inclusion of nonresearch articles did differ among the three journals, $\chi^2(2, N = 402) = 24.7$, $p < .01$, with *ResHE* reporting the highest percentage of research articles (88.2%) and *RevHE* presenting the greatest balance between research (62.5%) and nonresearch (36.5%) articles. Within the articles reporting original research, the most commonly used designs were correlational (56.3%), quantitative descriptive (17.0%), longitudinal (12.9%), case study (12.9%), and qualitative descriptive (9.4%). See Table 1 for relative frequencies of all designs reported across the three journals. Note that because designs were not mutually exclusive, some articles were included in multiple design categories. Chi-squared tests of independence were conducted to determine if the three journals differed in their reporting of the five most prominent designs. Tests were not conducted on the other designs due to small frequencies. To maintain a familywise error rate of .01 across the five tests, a Bonferroni adjusted alpha of .002 per test was used. The three journals differed in their relative reporting of studies using correla-

TABLE 1. Frequency of Designs Reported Across 402 Articles in the Three Journals

Design	Frequency	Percentage of Studies
Correlational	192	56.3
Descriptive–Quantitative	58	17.0
Longitudinal	44	12.9
Case Study	44	12.9
Descriptive–Qualitative	32	9.4
Ex Post Facto	13	3.8
Grounded Theory	12	3.5
Scale Development	7	2.1
Historical	5	1.5
Naturalistic Inquiry	4	1.2
True Experiment	4	1.2
Quasi-experimental	3	.9
Ethnography	3	.9
Meta Analysis	2	.6
Action Research	2	.6
Total Appearances	425[a]	

[a]A total of 425 different designs were used across the 402 articles.

tional, χ^2 (2, $N = 343$) = 58.4, $p < .002$, case study, χ^2 (2, $N = 343$) = 27.8, $p < .002$, and descriptive qualitative, χ^2 (2, $N = 343$) = 12.9, $p < .002$, designs. Correlational designs appeared in the majority (79.2%) of *ResHE* articles, whereas correlational designs were reported in only 28.6% of the *RevHE* studies. In contrast, both case study and descriptive qualitative designs were seen most often in *RevHE* and least often in *ResHE*. Table 2 presents frequencies of the five most commonly reported designs by journal.

Studies were also broadly categorized as "quantitative only," "qualitative only," and "mixed" in terms of their designs and methods. Studies coded as quantitative only were clearly the most frequently published (73.4%) across all three of the journals. Studies that were exclusively qualitative comprised the next most frequently published type of study (20.3%), with mixed designs used in only 6.3% of the studies. The three journals differed significantly in the relative proportions of broad design types, $\chi^2(4, N = 335) = 74.4$, $p < .01$, with *ResHE* reporting the greatest proportion of studies relying exclusively on quantitative methods (93.5%) and *RevHE* reporting the greatest proportion of exclusively qualitative studies (46.7%).

In terms of data collection methods/sources, most of the studies used surveys based on either primary (41.5%) or secondary (35.1%) data. Commonly used sources of secondary survey data included the *College Student Experiences Questionnaire* (*CSEQ*), *College Institutional Research Program* (CIRP) *Freshman Survey*, and data sets published by the National Center for Education Statistics (NCES), such as the *National Study of Postsecondary Faculty* and *High School and Beyond*. Thirty-one percent of the studies relied on the use of documents or artifacts (including reports, transcripts, and catalogues), and 23.8% used interviews. Frequencies for all data collection methods and sources are presented in Table 3. Among the three journals, *RevHE* reported significantly greater use of interviews, $\chi^2(2, N = 333) = 66.1$, $p < .002$, than either *JHE* or *ResHE*. No other differences were found among the journals with respect to

TABLE 2. Relative Frequency of Design Reported by Journal

	Journal		
Design	*JHE*	*RevHE*	*ResHE*
Correlational[a]	56	18	118
Descriptive–Quantitative	23	17	18
Longitudinal	12	10	22
Case Study[a]	13	23	8
Descriptive–Qualitative[a]	12	14	6

[a]Chi-square tests of independence were significant at $p < .002$.

TABLE 3. Frequency of Data Collection Methods and Sources Across the Journals

Data Source	Frequency	Percentage of Studies
Survey (primary)	136	41.5
Survey (secondary)	115	35.1
Documents	103	31.4
Interview	78	23.8
Observation	24	7.3
Test or Outcome Measure	8	2.4
Focus Group	8	2.4
Meta-Analysis	2	.6
Total Responses	474	

data source. The most frequently used data sources are reported for the three journals in Table 4.

Among the 252 studies reporting some type of quantitative analyses, the most popular inferential statistical procedure was multiple regression; either ordinary least squares (44.8%) or logistic (11.9%). In addition, 39.3% reported bivariate correlations, and 26.2% conducted either exploratory (22.2%) or confirmatory (4.0%) factor analysis. Although over 236 of the studies (78.6%) reported some type of quantitative descriptive information including means, standard deviations, frequencies ($n = 198$), or graphs/plots ($n = 38$), only 32 studies reported solely descriptive statistics. Relative frequencies of statistical procedures reported in at least 10 studies are listed in Table 5. Prevalence of the statistical techniques listed in Table 5 did not differ significantly across the three journals.

As mentioned earlier, statistical analyses were coded as basic, intermediate, or advanced for purposes of linking the methodological level of the literature with the level of the standard research training received by the bulk of he doc-

TABLE 4. Relative Frequency of Data Source Reported by Journal

	Journal		
Design	*JHE*	*RevHE*	*ResHE*
Survey (primary)	42	26	68
Survey (secondary)	33	21	61
Document	31	30	42
Interview[a]	30	40	8
Observation	12	9	3

[a]Chi-square tests of independence were significant at $p < .002$.

TABLE 5. Frequency of Statistical Analyses Reported Across the Journals

Statistical Analysis	Frequency	Percent of Studies
Descriptive Statistics[a]	198	78.6
Multiple Regression	113	44.8
Bivariate Correlation	99	39.3
Exploratory Factor Analysis	56	22.2
Oneway or Factorial ANOVA	55	21.8
Latent or Observed Variable Path Modeling	45	17.9
Chi-squared	44	17.5
Descriptive Graphs/Plots	38	15.1
Independent Samples T-test	34	13.5
Logistic Regression	30	11.9
Post Hoc Comparisons	29	11.5
MANOVA	16	6.3
Discriminant Analysis	10	4.0
Repeated Measures	10	4.0

[a]Descriptive statistics include means, standard deviations, and frequencies.

toral students. Interestingly, nearly half (48.9%) of the quantitative studies reported use of statistical procedures normally taught in advanced courses, while 33.5% and 17.6% reported the types of analyses typically taught in intermediate and basic courses, respectively. The three journals did differ significantly with respect to the level of analyses reported, χ^2 (4, $n = 272$) = 18.7, $p < .01$, with the journals reporting comparable proportions of analytic procedures at the intermediate level but different proportions at the basic and advanced levels. Of the 151 quantitative articles published by *ResHE*, over half (55%) used one or more statistical procedures that would be categorized as advanced, whereas only 35% and 44.4% of the quantitative studies published in *RevHE* and *JHE*, respectively, used advanced procedures. In contrast, over a third (35%) of the *RevHE* studies reported basic-level statistics as the highest level while only 9% of studies in *ResHE* used basic statistics as the highest level.

Regarding measurement qualities, 66% of the quantitative studies included clear operational definitions of the variables used in their analyses; however, only 31% presented thorough or even partial/past (9%) reliability estimates. Further, many of the quantitative studies used single-item indicators (43%) to measure the variables/constructs of interest, even when those constructs were clearly complex and multidimensional. Attention to measurement validity was even scarcer with only 2% reporting thorough validity information and another 15% reporting minimal validity support. Among the 235 studies using either primary or secondary survey data, only 38.6% addressed potential nonrespondent bias even though response rates were quite low in many cases.

CONCLUSIONS AND DISCUSSION

Our results support prior research (Goodwin and Goodwin, 1985), which has concluded that professionals require at least intermediate-level statistics to adequately comprehend most of the published research. However, our findings also suggest that more sophisticated procedures are beginning to appear in the higher education research literature, requiring knowledge of multivariate procedures including factor analysis, path modeling, and logistic regression. In comparing our results with those of Goodwin and Goodwin, we found that higher education professionals with training through intermediate-level statistics would only be comfortable reading about half of the journal articles we reviewed, although the proportions would differ somewhat among the three journals. The other half would require additional, more advanced training in statistics and research methods. Malaney (2002) has proposed that doctoral students training to be student affairs professionals be required to take a minimum of four research courses, which is consistent with our definition of the advanced methodology training needed to comprehend much of the published higher education literature. This differs from what Kuh, Bean, Bradley, Coomes, and Hunter (1986) found in the student affairs journals where the majority of articles were descriptive only, with less than a third utilizing any multivariate technique. Two observations are important here. First, our analysis included only general higher education journals, with student affairs journals excluded. Thus, the discrepancy in findings might be due to differences in content and focus between journals published in the student affairs area and those in general higher education. Kuh, Bean, Bradley, and Coomes (1986) and Kuh, Bean, Bradley, Coomes, and Hunter (1986) mentioned a similar disparity in their studies. A second important aspect to note is the time difference. Our study surveyed journals from 5 recent years (1996–2000) while the studies by Kuh and his associates examined journals that are now almost 2 decades old. Consequently, it is difficult to conclude if differences in methodologies are solely due to differences in content and focus or to changes that might have occurred naturally over time with greater sophistication (or perhaps some combination).

A limitation of the current study was the use of only descriptive and univariate, inferential statistical procedures due to the exploratory nature of the study. A future study might apply more sophisticated statistical procedures such as loglinear analysis or logistic regression to examine simultaneously a variety of factors that could account for use of particular methods and procedures. Such a study might examine author characteristics (such as gender, type of institution, and publication frequency), data source, and topic or content of the studies to determine if these factors are related to methodological characteristics of the studies.

One implication of our findings is that doctoral programs in higher education need to provide research training of sufficient depth and scope to prepare graduates not only to be proficient producers of research but also critical consumers of the published research. Given the preponderance of intermediate and advanced levels of analyses, as well as the diversity of designs and data collection methods reported in the three journals reviewed in our study, such training would likely require a sequence of four or more research methods courses. This recommendation, also offered by Malaney (2002), applies equally to higher education and student affairs practitioners as well as to those planning to pursue careers in academe. Though certainly the specific types of research skills necessary for different types of higher education positions might differ to some extent, all professionals in the field of higher education, whether at the master's or the doctoral level, should possess, at a minimum, the skills necessary to comprehend the published work in their field. Schroeder and Pike (2001) advocate that student affairs professionals need to obtain sufficient preparation in research methods so they can "understand the connection between research and practice" (p. 349), without which "those practitioners will never systematically apply knowledge to problems of significance" (p. 349).

As mentioned earlier, a trend we noted anecdotally was that a number of authors appeared across multiple articles. In many cases, these authors applied relatively sophisticated statistical procedures to large, complex secondary data sets. What is not known is from which institutions these authors received their doctoral training. Although beyond the scope of the present study, a future study might examine the frequency of authorship according to institution where doctoral degree was received. Such a study might reveal that the bulk of articles are coming from institutions with more rigorous research training requirements. Such a study might also reveal if in effect a limited number of higher education training programs are shaping the literature in the field.

Although our coding of measurement attributes was fairly limited, it was apparent from our findings that many researchers in higher education pay token attention to sound measurement principles. This was evident both in the wholesale absence of validity discussion as well as from the use of single-item measures of attitudes, perceptions, and so on, even when such measures were unwarranted by the complexity of the variables/constructs being studied. The lack of attention to measurement quality is disturbing given that even the most sophisticated statistical technique provides meaningless results if not performed on valid and reliable data. Other researchers have also expressed concern about inadequacies in measurement training and expertise both in the fields of psychology and education. In a study of methodological training among psychology majors, Aiken et al. (1990) found psychology students to be inadequately trained in basic psychological measurement. More recently, Thompson (2001) noted the

misuse and misinterpretation of reliability coefficients seen in many educational and psychological journals. Similarly, Kieffer et al. (2001) found only 50% of articles in *AERJ* and *JCP* correctly reported information on score reliability, which was somewhat higher than the 31% in our study. In the field of higher education, the inattention to measurement likely reflects a lack of appropriate measurement training as suggested by a survey of research requirements among higher education doctoral programs conducted by Lovell et al. (1999) and Lovell and Hutchinson (2003). Of the higher education programs responding to the survey, few required measurement courses, and most tended to require only introductory level, statistically focused courses. Thus, the lack of awareness about measurement issues in the three journals reviewed in the current study seems to mirror the general inattention to measurement in many doctoral training programs.

It was somewhat surprising to find in the current study the relatively limited inclusion of qualitative studies in the higher education literature. This differs somewhat from the finding by Elmore and Woehlke (1998) who found an increase in the use of qualitative methods over the past ten years. Although we found the *RevHE* included proportionally more qualitative studies than either quantitative or mixed design studies, the *JHE* and *ResHE* both reported mostly quantitative studies. More important than differences among the journals was the finding that relative prevalence of designs across the three journals was stable across the 5-year period addressed in our study. Perhaps 5 years is insufficient for revealing paradigmatic shifts or changes in publication patterns, despite the comparative instability of editorial boards. Or it might be that the field itself has stabilized in terms of the struggle between qualitative and quantitative paradigms. Nevertheless, it was not our intention to engage in the "paradigm wars" by quarreling over what should be in terms of relative frequency of methodological approaches in the three journals we reviewed, but only to chronicle what is.

Silverman (1987) posited, "perhaps journals both create and mirror their fields" (p. 40). If true, then the methodologies showcased in the three journals in our study suggest that higher education researchers possess fairly strong methodological skills in statistical analyses, but somewhat limited training in measurement. Moreover, the journals surveyed in the current study also model the expectation that doctoral students-in-training will need to become increasingly competent in a variety of research methods (including both qualitative and quantitative approaches) if they are to fully comprehend as well as successfully publish research in the three higher education journals reviewed.

Finally, faculty in higher education doctoral programs may want to consider whether graduate-level training in the field of higher education needs modifications based on the findings of the present study as well as on findings regarding

research requirements by Lovell et al. (1999) and Lovell and Hutchinson (2003). Not only might many doctoral graduates be ill-prepared to conduct research, as discussed previously, but many doctoral-level professionals may not possess the skills necessary to fully appreciate or even to understand the information presented about research methods and procedures in the published higher education literature. Findings from Kezar's (2000) recent study of higher education researchers' and practitioners' perceptions of the relevance and utility of higher education literature seem to support the possibility of inadequate research preparation. We are in no way suggesting that professional journals or the editorial boards that govern them should dictate graduate-level curriculum, but we are contending that those responsible for curriculum design need to be responsive to changes and methodological advances in the discipline, which are often disseminated through professional journals. The issue of research competency and training and its connection with the published research is one that warrants additional study as well as dialogue within the higher education community.

ACKNOWLEDGMENT

An earlier version of this paper was presented at the annual meeting of the American Educational Research Association, Montreal, 1999.

APPENDIX: Coding Sheet for Higher Education Journals

Journal Name:
JHE_____ RHE_____ Res.HE_____

Journal Year: _____

Article Citation:

_____

_____

Type of Article:
Expository/lit review/anecdotal _____
Expository supported by data _____
Opinion/position paper _____
Qualitative Study _____
Quantitative Empirical _____

Topic: _____

APPENDIX: (*Continued*)

Design(s) Used:

True experiment: _____
Quasi-experiment _____
Correlational _____
Ex Post Facto or Causal Comparative _____
Descriptive (quantitative) _____
Descriptive (qualitative) _____
Developmental/longitudinal _____
Ethnography _____
Case Study _____
Historical _____
Action research _____
Meta-analysis _____
Test/scale validation _____
Other _____

Broad Design Type:

Quantitative only _____
Qualitative only _____
Mixed _____

Data Collection Method/Source:

Survey (primary) _____
Survey (secondary) _____
Observation _____
Interview _____
Documents/artifacts _____
Meta-analysis _____
Test or attitude outcome measure (experiment) _____
Other _____

Types of Statistical Analysis Used

Basic

ANOVA oneway _____
Bivariate correlations _____
Chi-squared _____
Descriptives (mean, sd, %) _____
t test (indep. samples) _____
t test (paired) _____

APPENDIX: (*Continued*)

Intermediate

ANCOVA (oneway) _____
ANCOVA (factorial) _____
ANOVA factorial _____
Multiple regression _____
Path modeling (observed var.)
Planned orthogonal comparisons _____
Post hoc comparisons _____

Advanced

Cluster analysis _____
Discriminant analysis _____
Factor or principal components analysis _____
Logistic regression _____
Logit/loglinear _____
MANOVA/MANCOVA
Repeated Measures _____
Structural equation modeling _____
Other _____

Descriptives *only* _____

means, sd _____
%, proportions, frequencies _____
graphs, charts, figures _____

Highest Level Statistics Used

Basic _____ Intermediate _____ Advanced _____

Quality of Article:

Defines population and/or sampling frame:
No ____ Population ____ Sampling frame ____
Provides discussion of statistical assumptions:
No ____ Yes, minimal ____ Yes, thorough ____
Provides clear operational definition of variables:
No ____ Yes, partial ____ Yes ____
Presents scale reliability:
No scale used ____ No ____ Yes, partial or past ____ Yes ____
Presents validity info:
No ____ Yes, minimal ____ Yes, thorough ____
For surveys, addresses potential nonrespondent bias:
No ____ Yes ____
Other flaws noted: _____

REFERENCES

Aiken, Leona S., West, Stephen G., Sechrest, Lee, and Reno, Raymond R. (1990). Graduate training in statistics, methodology, and measurement in psychology. *American Psychologist* **45**(6)**:** 721–734.

Baumberger, Julie P., and Bangert, Arthur W. (1996). Research designs and statistical techniques used in the *Journal of Learning Disabilities* 1989–1993. *Journal of Learning Disabilities* **29**(6)**:** 313–316.

Blimling, Gregory S. (ed.) (2001). Scholarship in student affairs reconsidered [Special issue]. *Journal of College of Student Development* **42**(4).

Cabell, David W. E., and English, Deborah L. (eds.) (2002a). *Cabell's Directory of Publishing Opportunities in Educational Curriculum and Methods* (6th Ed., Vol. 2), Cabell Publishing, Beaumont, TX.

Cabell, David W. E., and English, Deborah L. (eds.) (2002b). *Cabell's Directory of Publishing Opportunities in Educational Psychology and Administration* (6th Ed., Vol. 2), Cabell Publishing, Beaumont, TX.

Carpenter, Stanley (2001). Student affairs scholarship (re?)considered: Toward a scholarship of practice. *Journal of College Student Development* **42**(4)**:** 301–318.

Cook, Thomas D., and Campbell, Donald T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Dallas.

Creswell, John W. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, Merrill Prentice Hall, Columbus, OH.

Elmore, Patricia B., and Woehlke, Paula L. (1998, April). Twenty years of research methods employed in American Educational Research Journal, Educational Researcher, and Review of Educational Research. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Gall, Meredith D., Gall, Joyce P., and Borg, Walter R. (2003). *Educational Research: An Introduction* (7th Ed.), Allyn & Bacon, Boston.

Goodwin, Laura D., and Goodwin, William L. (1985). Statistical techniques in *AERJ* articles, 1979–1983: The preparation of graduate students to read the educational research literature. *Educational Researcher* **14**(2)**:** 5–11.

Hummel, Linda J. (1984). An approach to a study of physician learning patterns. *MOBIUS* **4**(4)**:** 48–50.

Isaac, Stephen, and Michael, William B. (1984). *Handbook in Research and Evaluation for Education and the Behavioral Sciences* (2nd Ed.), EdITS, San Diego.

Kezar, Adrianna (2000). Higher education research at the millennium: Still trees without fruit? *Review of Higher Education* **23**(4)**:** 443–468.

Kieffer, Kevin M., Reese, Robert J., and Thompson, Bruce (2001). Statistical techniques employed in *AERJ* and *JCP* articles from 1988 to 1997: A methodological review. *Journal of Experimental Education* **69**(3)**:** 280–309.

Krathwohl, David R. (1998). *Methods of Educational and Social Science Research: An Integrated Approach* (2nd Ed.), Longman, New York.

Kuh, George D., Bean, John P., Bradley, Russell K., and Coomes, Michael D. (1986). Contributions of student affairs journals to the literature on college students. *Journal of College Student Personnel* **27**(4)**:** 292–304.

Kuh, George D., Bean, John P., Bradley, Russell K., Coomes, Michael D., and Hunter, D. E. (1986). Changes in research on college students published in selected journals between 1969 and 1983. *Review of Higher Education* **9**(2)**:** 177–192.

Love, Christine (1996). How nurse teachers keep up-to-date: Their methods and practices. *Nurse Education Today* **16**(4)**:** 287–295.

Lovell, Cheryl D., and Hutchinson, Susan R. (2003). *Research training requirements in higher education doctoral programs*. Manuscript in preparation.

Lovell, Cheryl D., Hutchinson, Susan R., and Fairweather, James R. (1999, November). *Graduate student research preparation in higher education: Too little, just right, or too much . . . Implications for our future*. Paper presented as the Annual Meeting of the Association for the Study of Higher Education, Miami, FL.

Malaney, Gary D. (2002). Scholarship in student affairs through teaching and research. *NASPA Journal* **39**(2)**:** 132–146.

Milam, John H. (1991). The presence of paradigms in the core higher education journal literature. *Research in Higher Education* **32**(6)**:** 651–668.

Roper, Larry D. (ed.) (2002). The scholarship of student affairs [Special issue]. *NASPA Journal* **39**(2).

Schroeder, Charles C., and Pike, Gary R. (2001). The scholarship of application in student affairs. *Journal of College Student Development* **42**(4)**:** 342–355.

Silverman, Robert J. (1987). How we know what we know: A study of higher education journal articles. *Review of Higher Education* **11**(1)**:** 39–59.

Thompson, Bruce (2001). Significance, effect sizes, stepwise methods, and other issues: Strong arguments move the field. *The Journal of Experimental Education* **70**(1)**:** 80–93.

Volkwein, J. Fredericks, Carbone, David A., and Volkwein, Edward A. (1988). *Research in Higher Education*: Fifteen years of scholarship. *Research in Higher Education* **28**(3)**:** 271–280.

Received January 17, 2003.