2013

A Meta-View of Multivariate Statistical Inference Methods in European Psychology Journals

Lisa L. Harlow
University of Rhode Island, lharlow@uri.edu

Elly Korendijk

See next page for additional authors

Citation/Publisher Attribution
Harlow, L. L., Korendijk, E., Hamaker, E. L., Hox, J., & Duerr, S. R. (2013). A meta-view of multivariate statistical inference methods in European psychology journals. *Multivariate Behavioral Research*, *48*, 749-774. doi: 10.1080/00273171.2013.822784
Available at: http://www.dx.doi.org/10.1080/00273171.2013.822784

Authors
Lisa L. Harlow, Elly Korendijk, Ellen L. Hamaker, Joop Hox, and Sunny R. Duerr

2nd Final Revision and Preprint: May 23, 2013 for *Multivariate Behavioral Research* article:

Harlow, L. L., Korendijk, E., Hamaker, E. L., Hox, J., & Duerr, S. R. (2013). A meta-view of multivariate statistical inference methods in European psychology journals. *Multivariate Behavioral Research, 48*, 749-774.

A Meta-View of Multivariate Statistical Inference Methods in European Psychology Journals

Lisa L. Harlow, University of Rhode Island

Elly Korendijk, Ellen L. Hamaker and Joop Hox, Utrecht University

Sunny R. Duerr, University of Rhode Island

Author Note

Lisa L. Harlow, Department of Psychology, University of Rhode Island; Elly Korendijk, Department of Methodology and Statistics, Utrecht University (Note: E. Korendijk is now at Biometris, Wageningen University and Research Center); Ellen L. Hamaker, Department of Methodology and Statistics, Utrecht University; Joop Hox, Department of Methodology and Statistics, Utrecht University; and Sunny R. Duerr, Department of Psychology, University of Rhode Island (now at Department of Secondary Education, State University of New York at New Paltz).

This research was supported in part by Grant 0720063 from the National Science Foundation to the first author, and funding from the Dean's Office and the Provost's Office at the University of Rhode Island. Thanks are also extended to staff, faculty and students at the Department of Psychology and the Cancer Prevention Research Center at the University of Rhode Island, and at the Department of Methodology and Statistics, Utrecht University. Correspondence concerning this article should be addressed to Lisa L. Harlow, Department of Psychology, 10 Chafee Rd, Suite 8, University of Rhode Island, Kingston, Rhode Island, USA 02881-0808. E-mail: lharlow@uri.edu

A Meta-View of Multivariate Statistical Inference Methods in European Psychology Journals

Abstract

(150 words)

We investigated the extent and nature of multivariate statistical inferential procedures used in eight European psychology journals covering a range of content (i.e., clinical, social, health, personality, organizational, developmental, educational, and cognitive). Multivariate methods included those found in popular texts that focused on prediction, group difference, and advanced modeling: multiple regression, logistic regression, analysis of covariance, multivariate analysis of variance, factor- or principal components analysis, structural equation modeling, multilevel modeling, and other methods. Results revealed that an average of 57% of the articles from these eight journals involved multivariate analyses with a third using multiple regression, 17% using structural modeling, and the remaining methods collectively comprising about 50% of the analyses. The most frequently occurring inferential procedures involved prediction weights, dichotomous *p*-values, figures with data, and significance tests; with very few articles involving confidence intervals, statistical mediation, longitudinal analyses, power analysis, or meta-analysis. Contributions, limitations and future directions are discussed.

Keywords: multivariate analyses, statistical inference, effect sizes, significance tests

Multivariate Statistical Inference Methods

Making accurate statistical inferences is important in all fields of research, with a growing call for reform (e.g., Harlow, Mulaik & Steiger, 1997; Kline, 2004). Emphasis is moving to a reconsidering or at least supplementing of null hypothesis significance testing (NHST) and dichotomous probability (*p*) values (Cumming, 2008) with other methods including effect sizes (APA, 2010; Cohen, 1994; Kelley & Preacher, 2012; Kirk, 1996; Peng, Chen, Chiang & Chiang, 2013), confidence intervals (APA, 2010; Cumming, 2012; Lai & Kelley, 2011), figures and graphics (Cleveland, 1993; Friendly, 2000), meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2009; Cafri, Kromrey, & Brannick, 2010), statistical power (Aberson, 2010; Cohen, 1962, 1988; Sun, Pan & Wang, 2011), and mediation (Baron & Kenny, 1986; MacKinnon, 2008), among others. Moreover, although there is some literature on multivariate methods regarding inferences (e.g., Bodnar, Bodnar, & Gupta, 2010), effect sizes (e.g., Hess, Hogarty, Ferron, & Kromrey, 2007; Raudenbush, Becker, & Kalaian, 1988; Steyn & Ellis, 2009) and meta-analysis (e.g., Becker, 2000; Chan, & Arvey, 2012; Nam, Mengersen, & Garthwaite, 2003), statistical inference regarding multivariate methods has received scarce attention. In particular, although there have been studies of univariate effect sizes and inferential procedures (e.g., Cohen, 1962; Cumming, 2012; Kirk, 1996; Rossi, 1990), little is known about the extent and nature of multivariate inferential methods in substantive research journals. Moreover, although there is growing interest in international activities in psychology (e.g., See the American Psychological Association Office of International Affairs: http://www.apa.org/international/) and the use of statistical methods in Europe (e.g., See the European Association of Methodology: http://www.eam-online.org/), there has not yet been an investigation of the use of multivariate inferential methods in European journals. Getting a view of the landscape in these journals can offer greater awareness of what is currently being conducted in existing literature, and may offer insight into what could be done to improve the nature of multivariate inference in the future.

Purpose of the Study

The current paper describes a survey of major European psychology journals that were specifically selected so as to cover a wide spectrum of content areas (e.g., clinical, personality, health, social, cognitive, applied, developmental, and educational) to investigate multivariate methods and inferential procedures in psychological research. The main goal was to gain a meta-view of how multivariate statistical inferences were made and reported in the psychological literature in European journals. In particular, it was of interest to see: how much multivariate analyses were used in each of the journals examined, what kinds of methods were used, and how much of the multivariate research emphasized traditional NHST procedures such as significance tests and dichotomous *p*-values, as well as more informative methods such as ESs, CIs and other inferential procedures. Further, we wanted to explore the relationship between the various multivariate methods used and the specific inferential procedures that were used to gage current practices and suggest future improvements.

Multivariate Methods Assessed

For our survey, we considered a method as "multivariate" if it analyzed data from multiple variables or time points, or provided an extension from a univariate or bivariate analysis. The selection of multivariate methods is typical of what is covered in popular multivariate texts (e.g., Grimm & Yarnold, 1995, 2000; Stevens, 2009; Tabachnick & Fidell, 2012).

Multivariate methods examined include at least eight methods. These involved two group difference methods of analysis of covariance (ANCOVA), and multivariate analysis of variance (MANOVA) that allow for inclusion of one or more covariates or multiple dependent variables, respectively. At least three prediction methods were also examined, including multiple regression (MR), logistic regression (LR), and multilevel modeling (MLM). MR is useful for predicting a continuous outcome and LR is used with a categorical, usually dichotomous, outcome. MLM takes into account naturally occurring groups when examining prediction. We also included structural methods of factor- or principal components analysis (FA or PCA) that show how a single set of measures relates to underlying factors or components; and structural equation modeling (SEM), which allows predictions among latent factors that are each indicated by one or more measured variables. Other less used multivariate methods such as canonical correlation, cluster analysis, discriminant analysis, dynamic factor analysis, latent class analysis or latent transition analysis, generalized estimating equations (GEE), multidimensional scaling, survival analysis or hazard analysis, and time series were coded under the heading of "Other."

Inferential Procedures Assessed

It addition to the multivariate methods that are being used, we were also interested in the use of specific kinds of inferential procedures, which are briefly described, below.

Significance Tests

Historically, the procedure known as NHST has been conducted to determine whether results are significantly different from chance (APA, 2010). Although NHST has received considerable critical assessment (e.g., Balluerka, Gómez, & Hidalgo, 2005; Cohen, 1994; Cumming 2012; Schmidt & Hunter, 1997), it also has proponents (e.g., Chow, 1996; Mulaik, Raju & Harshman, 1997) and continues to be a popular procedure in social science research. In the current study, it was of interest to examine how much significance tests (e.g., F, χ^2) were used in multivariate research in European journals.

Probability Values

Probability values (i.e., *p*-values) provide information on the chance of finding a result as different, or more so, than that found in a research study, if the null hypothesis is actually true. Critics of *p*-values (e.g., Cumming, 2008) question the tendency to make a dichotomous decision (e.g., reject the null hypothesis, if, say, $p < .05$). The use of exact *p*-values (e.g., $p = .02$) is viewed more favorably (e.g., McGrath, 2011) as it provides more information and avoids limiting interpretation to an either/or decision rule. Cumming (2008, 2012), however, claims that *p*-values are highly unstable and do not replicate well. Thus, we were interested in examining how much *p*-values, whether dichotomous or exact, were used in multivariate research.

Proportion (or %) of Shared Variance between Variables

The proportion of shared variance between two sets of variables is a useful effect size (ES), particularly for multivariate analyses. We investigated whether proportions or percentages of shared variance indices (e.g., η^2, R^2, and ω^2 or percentage of variance in items explained by factors or components) were provided in European journals, as this would be informative to researchers interested in an over-arching or macro-level ES from a multivariate analysis.

Weights (e.g., Loadings, B,β & Odds Ratios).

A more micro-level form of ES that is often used in multivariate research is a weight that shows how much a single variable contributes to an analysis (Harlow, 2005). Weights can involve loadings, odds ratios, or regression weights, for example. For our study, it was of interest how much weights were used in the European journals surveyed in this study, revealing the extent of providing micro-level ES coefficients in a correlational or predictive framework.

Figures with Data

There is a long history of visually depicting data to illuminate the main trends or patterns. Tukey (1977) championed the idea of exploratory data analysis that visually illuminates the data.

Pearson (1885) is credited with introducing the histogram to indicate the frequencies of different levels of a single quantitative variable. Cleveland (1984, 1993) proposed the dot plot as an alternative to a pie chart or histogram. We investigated how often figures were used for conveying results from multivariate analyses. We also use several figures for highlighting the results from our survey of the European journal articles.

Standard Errors (SEs)

Standard errors provide an indication of how spread out scores are and how precise a parameter estimate is. We explored whether standard errors were reported for the multivariate analyses used in these articles.

Confidence Intervals (CIs)

CIs provide an indication of the range of values expected in the population for a specific sample parameter estimate. The use of CIs has been advocated by many (e.g., Cumming, 2008, 2012; Mendoza, & Stafford, 2001; Steiger & Fouladi, 1992; Thompson, 2007) to indicate the degree of precision, or conversely the degree of uncertainty about parameter estimates and ESs. We examined how much CIs were reported when presenting multivariate analyses in the European journal articles.

Fit and Other Indices

Fit indices are sometimes presented to assess how well data match an expected model and can be a useful supplement to multivariate significance tests and *p*-values, particularly as the latter are more likely to be dependent on sample size than fit indices (Bentler & Bonett, 1980). For example, SEM analyses tend to give values for fit indices, such as the comparative fit index (CFI: Bentler, 1990), the root mean square error of approximation (RMSEA: Steiger & Lind, 1980); or the Akaike information criterion (AIC: Akaike, 1987), among others. In multivariate group difference methods such as ANCOVA and MANOVA, where eta-squared or omega-squared is often the multivariate effect size of interest, Cohen's d (Cohen, 1988) that is used to present a univariate ES for a mean could be viewed as a supplemental fit index or effect size. In our study, the use of supplemental fit indices and effect sizes was also investigated.

Comparing Multiple Statistical Models

Recent literature calls for more statistical modeling (e.g., Harlow, 2010; Rodgers, 2010), and comparing multiple models to assess goodness of fit (e.g., Burnham & Anderson, 2002; Maxwell & Delaney, 2004). In the current paper we examined whether multiple models were tested and assessed in the articles that used multivariate analyses in the European journals we reviewed.

Mediation (and Moderation)

Statistical mediation and moderation (e.g., Baron & Kenny, 1986; MacKinnon, 2008) involve consideration of intervening or potentially confounding variables, respectively, to consider when examining relationships between independent and dependent variables. It was of interest to see whether either mediation or moderation was used in the multivariate analyses reported in the journals we surveyed.

Longitudinal Data

Longitudinal data involves assessing variables across multiple time points to yield stronger scientific inferences regarding causal patterns (e.g., Singer & Willett, 2003; Skrondal & Rabe-Hesketh, 2004). We investigated whether the multivariate analyses reported in the eight European journals used longitudinal data.

Other Considerations (e.g., Meta-analysis, Power analysis)

Other statistical inference procedures were considered. Meta-analysis (e.g., Glass, 1976; Hunter & Schmidt, 2004) is a useful method of culling across multiple studies to arrive at a more precise parameter estimate or ES. Statistical power is the probability of correctly rejecting the null hypothesis. The use of statistical power analysis is linked primarily with Cohen (1988) who advocated finding the needed sample size to find an expected ES with a reasonable (e.g., 80%) level of power. We explored whether meta-analysis or power considerations (e.g., Cafri, Kromrey, & Brannick, 2010; Rossi, 1990) were discussed in the multivariate articles in these European journals.

Questions Assessing Multivariate Statistical Inferences

For our study, we investigated several questions regarding the nature of multivariate methods and multivariate statistical inference, as well as the links between the methods and inferential procedures:

1. First, do major European psychology journals vary as to the number of multivariate analyses published?
2. Second, what are the kinds of multivariate methods that are used in these European psychology journals?
3. Third, what kinds of inferential procedures are used with the multivariate methods?
    a. Are traditional significance tests used to assess whether multivariate results are statistically different from chance and if so what kinds of *p*-values are being used: e.g., dichotomous: $p < .05$; or exact: $p = .03$?
    b. Are macro-level shared variance effect sizes used to gauge the magnitude of the overall multivariate effect, and are there any specific, micro-level effects, such as weights or loadings, which are examined?
    c. Are CIs used to indicate the degree of uncertainty around statistical effects and if so, which effects tend to use CIs?

d. Are there other inferential or similar research methods (e.g., visually displaying data in figures, or using longitudinal data, mediation or moderation, multiple models tested, meta-analyses, power analyses, etc.) to help provide insight and inferences about the findings?

4. Fourth, what was the pattern of multivariate method use that was prominent in each of the eight European psychology journals?
5. Finally, what is the overarching meta-view as to what kinds of inferential procedures tend to be used with each of the multivariate methods in these journals? This analysis will allow a more informed perspective on what are the current methodological practices in these journals and provide a basis for suggesting how the field might proceed from here to improve the nature of multivariate inferences in the future.

To address these questions, a survey of European journals in the field of psychology was conducted. The sample and measures used, as well as the coding procedures, are described below in the Methods section. We subsequently present Results and Discussion that address the five questions with a set of distributional descriptions, a dot plot, and a nonlinear canonical correlation analysis to reveal links between the use of multivariate methods and inferential procedures in the journals surveyed. A Contributions and Benefits section summarizes findings from the five questions, plus Limitations, Future Directions and a brief Conclusions section.

Method

Sample

All of the articles published in 2008 for eight major European journals were examined that covered a stratified array of selected substantive areas. The choice of journals was made after discussions with several European multivariate researchers who suggested a variety of content areas and specific journals that addressed these content areas. Across the selected eight journals, there were 456 articles – excluding editorials, errata or comments – 259 of which used multivariate analyses. Thus, for this sample of journals, an average of 57% of the articles used some form of multivariate analysis as described previously. The unit of analysis is the multivariate method, providing a sample of 365 analyses that were conducted in these 259 articles that used multivariate methods, as a number of the articles used more than one different kind of multivariate analysis. The eight journals are listed in Table 1, along with their impact factor, the number of multivariate articles, and the total number of research articles published in each journal (with an average of 57 articles published per journal in 2008). At the bottom of Table 1, notice that impact factor is positively correlated with the number of multivariate articles.

----Insert Table 1 about here----

Measures

The "measures" that were examined in this study include information about use of any of the multivariate method analyses mentioned (i.e., ANCOVA, MANOVA, MR, LR, MLM, FA-PCA, SEM, Other), and the specific types of inference procedures used: Significance tests (e.g., F, χ^2, Likelihood Ratio); probability values (e.g., $p < .05$, $p = .04$), proportion (or %) of shared variance between variables (e.g., R^2, η^2); weights (e.g., loadings, B,β, Odds Ratios); figures with data; standard errors (of parameter estimates); CIs; additional fit indices, and univariate effect sizes or significance tests. (e.g., CFI, RMSEA, AIC, GFI, Cohen's *d*, *t*-test); statistical mediation or moderation; and other considerations (e.g., comparing multiple statistical models, meta-analysis, and statistical power analyses).

The main focus was on investigating the various types of multivariate methods that were conducted, the specific kinds of inferential procedures that were used, and the links between these multivariate methods and inferential procedures in a broad sampling of European psychology journals.

Coding Procedure

Coders, all of whom were very familiar with multivariate methodology, were drawn from the co-authors. To establish thorough, accurate, and reliable coding of articles, coders were first trained to use a coding sheet designed to identify the relevant data in an organized fashion; the coding sheet for this project can be seen in the Appendix. This training included a discussion about what types of analyses should be considered multivariate in nature and what information should be captured for each type of analysis. Coders downloaded electronic copies of all articles in 2008 from the eight journals, and manually reviewed each electronic copy to assess whether multivariate methods or inferential procedures were used. All 456 articles that comprise the 2008 issues of the eight selected European Journals were read and reviewed for relevance.

After all of the coding was completed, the first author went through the entire set of data to verify that coding was accurately and consistently recorded across coders. The main change that was needed was to make sure that any analysis that involved structural equation modeling was recorded in that category, with any additional information also recorded. For example, latent growth modeling (LGM) was sometimes coded as SEM and sometimes listed as “Other” under type of method. We revised the data to list LGM as an SEM method, and also checked “longitudinal data” under the various inferential procedure variables.

For the purpose of this project, each article contained three levels of information (Harlow, 2005). The first level of information is considered the macro-level and asks, “Does the article contain any multivariate analyses?” The second level of information is considered the mid-level and is embodied by answering the question, “What types of multivariate analyses are included in the article?” The third level of information is the micro-level and is embodied by answering the question, “What specific statistical information is provided in the article?” Equipped with these questions and the coding sheet, the coders proceeded to review each article.

The first task for each article was to determine whether or not the article contained any multivariate analyses. This was commonly obtained by reviewing the abstract, methods, and results sections of the article, or using the search function within Acrobat Reader, to find some key terms commonly used when conducting multivariate analyses, such as *regression*, *multivariate*, and *structural modeling* (to name a few). Additionally, the mention of certain types of computer software could also indicate the existence of multivariate analyses (i.e., AMOS, Mplus, or LISREL for structural modeling methods, or MetaWin for meta-analysis methods). Finally, the existence of path diagrams was often indicative of path analysis or structural equation modeling techniques. If the article contained any type of multivariate analysis, it was to be included in the study and the article’s key points of information were recorded (e.g., journal name, volume, and issue number; first author’s last name; page numbers for articles, etc.).

Once the article was deemed to contain multivariate analyses, the coder would search for the specific types of analyses included. If an article had multiple different multivariate analyses, each type of analysis was recorded. Along with each type of analysis, the coder was looking for specific micro-level details regarding the analysis, such as the type of omnibus statistic used, how *p*-values were reported (dichotomous or exact), whether CIs and/or ESs were reported, how data was visually represented (graphs with or without data and/or error bars), and other information pertinent to the accurate reporting of multivariate statistics. Micro-level detail was recorded for each different type of multivariate analysis separately (i.e., a multiple regression analysis could have an omnibus *F*-test, and a logistic regression analysis might have a χ^2; should both of these types of analysis show up in the same article, both entries would be listed individually, noting each analysis and its omnibus test on the coding sheet). As a final step, any additional information from the article that was not specified was entered as a note on a coding sheet; such notes may include specific tests (i.e., Baron & Kenny test for mediation) or types of analyses that were not included in the typical analyses list (such as genetic modeling).

The final data set was a file that included 365 rows, one for each multivariate analysis conducted across the 259 articles, which each contained 1 to 4 multivariate analyses. The columns in the file included information about the journals that specified the multivariate methods and inferential procedures used in the analyses. The final coded file was used for subsequent analyses, using SPSS, to assess the nature and extent of multivariate methods and inferential procedures used in these European psychology journals.

Results and Discussion

Several kinds of analyses were conducted to summarize and provide a meta-view of the data from the multivariate analyses coded from the eight European journals. Results address each of the five questions asked earlier (i.e., 1. Is multivariate method use different across journals? 2. How much are each of the multivariate methods used? 3. How much are each of the inferential procedures used? 4. What is the pattern of multivariate method use across the eight journals? and 5. Can the data regarding multivariate methods and inferential procedures be examined from a broad and encompassing perspective to provide insight on the patterns of current research analyses in European psychology journals?). Descriptive and multivariate data analyses were conducted to examine the multifaceted nature of the data regarding multivariate methods and inferential procedures used. Findings are then summarized with suggestions for future directions.

Three sets of columns in Table 2 show the micro-level distributional pattern of multivariate analyses for one kind of variable that address the first three questions regarding: particular European psychology journal, type of multivariate method, and kind of inferential procedure, respectively. Figure 1 addresses question 4 about multivariate method use across the journals. Table 3 and Figure 2 highlight the intersection of multivariate methods and inferential procedures used to address question 5.

Multivariate Articles Reported in each European Psychology Journal

Addressing question 1, the first set of columns of Table 2 shows the percentages of multivariate analyses across the eight European psychology journals, where the Journal of Child Psychology and Psychiatry, a prominent European clinical psychology journal, used the most multivariate analyses of all the journals examined. The European Journal of Social Psychology and Psychology and Health journals followed; on down to the two journals with the least percentage of use -- the European Journal of Psychology of Education and the European Journal of Cognitive Psychology. Thus, for this selection of journals, there were clear differences in the amount of multivariate analyses used.

----Insert Table 2 about here----

Percentage of Analyses Conducted per Multivariate Method

To address the second question, the second set of columns of Table 2 shows the percentage of analyses conducted for each of the multivariate methods surveyed across the journals. The method with the largest use was multiple regression, followed by structural equation modeling, then LR, and FA-PCA, on down to the least use reported for MLM. Here again, there appear to be differences as to which multivariate methods are most widely used in these journals.

Percentage and Kinds of Inferential Procedures Used

The third question was examined with the percentages of use for 14 inferential procedures given in the third set of columns of Table 2. The pattern reveals that micro-level ES predictive weights or loadings are reported in almost two thirds of the analyses in these journals; followed by at least 50% reported use for dichotomous *p*-values, figures, and significance tests. There was also moderate use of macro-level shared variance ESs, fit indices, and comparing multiple models. The use of exact *p*-values, standard errors, CIs, and mediation or moderation analyses was low; and longitudinal analyses, meta-analyses and statistical power analyses were reportedly used very rarely in these journal issues. Note that the sum of the percentages exceeds 100% as most analyses involved the use of several inferential procedures. The pattern of percentages also reveals that micro-level effects sizes (e.g., regression weights and factor loadings) are provided quite often, which is encouraging to see, although traditional significance tests and dichotomous *p*-values are used almost as much, which is not consistent with recent reform guidelines.

Number of Analyses Conducted per Multivariate Method for each Journal

To answer the fourth question, we examined whether the use of specific multivariate methods varied across the selection of European psychology journals examined in this study. A chi-square test of independence was conducted, revealing a significant relationship between frequency of the eight specific multivariate methods used and the eight specific journals: χ^2 (49, N = 365) = 93.60, $p < .001$, $\varphi = .51$. Thus, there is .26 (i.e., the square of .51) shared variance (95% CI: [.154, .311] using the Steiger & Fouladi, 1992 R2 program) between the methods used and the specific journal, revealing a moderately large effect size. Figure 2 presents a dot plot visually depicting this relationship. The number of analyses conducted for each type of multivariate method is listed along the horizontal X-axis, for each of the eight European journals with the 8 methods repeated for each journal along the vertical Y-axis. This type of display is useful for integrating several facets of data in a single coherent graph. Cleveland (1984, 1993) developed the dot plot after conducting cognitive experiments that revealed that people process the data more accurately and easily when following across a series of dots indicating the frequencies or values of the outcome variable of interest. This goal seems to be met in Figure 2 where it is relatively easy to discern that the cognitive journal had the least amount of multivariate use (with MR mostly used), and the educational and developmental journals each having slightly more multivariate use (again with more MR and little use of the other methods), on up through the organizational, personality, health, social and then the clinical journals showing increasing multivariate use, again including a predominance of MR, as well as SEM, and some other kinds of analyses. It is interesting to note that MR use was higher than other multivariate method use across all eight journals, whereas the least used multivariate method varied across the journals with MLM less endorsed most often, and LR also turning up among the least used methods.

----Insert Figure 2 about here----

Macro-Level Multivariate Analysis of the Data

To address our fifth question about exploring a broad-based view of multivariate methods and inferential procedures used, we conducted a nonlinear canonical correlation analysis, sometimes called correspondence analysis (CA) (Greenacre & Hastie, 1987) or OVERALS (Van de Geer, 1987). Although a conventional canonical correlation analysis can be conducted with discrete variables that are dummy-variable coded such as are used in the current study (Cohen, Cohen, West, & Aiken, 2003; Maxwell, 1961), it is preferable to consider a procedure that specifically allows categorical variables. CA appears appropriate (kindly suggested by one of the anonymous reviewers on a previous draft) and provides an overall procedure for analyzing the correspondence among two or more sets of categorical variables using an alternating least squares (ALS) method. The nonlinear CA procedure (e.g., Yazici, Öğüş, Ankarali, & Gürbüz, 2010) is a form of optimal scaling that is similar to conventional canonical correlation except that it allows for the use of categorical variables such as those used to code the multivariate methods and inferential procedures used in the European journals. The CA procedure summarizes the data to present the most homogeneous and lowest number of dimensions needed to depict the relationships among the sets of variables.

In the current study, CA was used to assess whether two dimensions could adequately summarize the data given the two sets of variables (i.e., the set of multivariate methods and the set of inferential procedures). The two eigenvalues for the set of 8 multivariate methods and the set of 14 inferential procedures were .884 and .855, respectively, for the two dimensions. The "fit" in a CA is simply the sum of the eigenvalues, which equals 1.739 in this case. Dividing each eigenvalue by the summed fit value, provides an indication that the first dimension explained 50.8% of the available variation in the CA, and the second dimension explained slightly less, 49.2%. Together, the two dimensions explained 87% (i.e., [.884 + .855]/ 2 dimensions) of the variation in what are called the object scores for the variables. Object scores are values for each "subject" or "object" that in this case represent the 365 analyses from the European journals, where the scores quantify the categories of the sets of measures (i.e., multivariate methods and inferential procedures) for each analysis (see Yazici et al, 2010).

Component loadings can be obtained (See Table 3 where the most salient loadings are bolded) that reveal the correlations between the object scores and the actual data that have been optimally scaled in the CA. Figure 2 provides a visual depiction of the component loadings, showing how closely related are the eight multivariate methods (i.e., ANCOVA, MANOVA, LR, MR, MLM, FA-PCA, SEM, Other) with the 14 specific inferential procedures (i.e., weights, dichotomous *p*-values, figures, significance tests, percentage of shared variance, supplemental fit index use, multiple models, exact *p*-values, standard errors, CIs, mediation-moderation, longitudinal design, meta-analysis and power analysis). Variables that are plotted far from the intersection of the horizontal and vertical dimension axes indicate high loadings, with variables in similar proximity having more relationship. This information is useful in revealing the patterns of multivariate methods and inferential procedures that emerge in the literature for the eight European psychology journals.

----Insert Figure 2 about here----

From Table 3 and Figure 2, several patterns emerge. First, FA-PCA and percentage of variance explained have rather high (negative) loadings on dimension 1 and are relatively close in proximity. This pattern suggests that factor- and principal component analyses tend to report values indicating the percentage of variance in the variables that is explained by the factors or components. Further, significance tests, dichotomous *p*-values, and structural equation modeling appear to have moderately high positive loadings on dimension 1, suggesting that these three occur together in analyses. Thus, the first dimension appears to summarize the use of structural equation modeling with probability values and chi-square tests of model fit, and contrasts this with the use of factor or principal component analysis and percentages of explained variance.

Second, it appears that the variables indicating the use of weights and multiple regression are both loading rather highly on the second dimension (see upper portion of Figure 2); and that the standard errors variable (see Table 3 and upper middle section of Figure 2) loads moderately on dimension 2, although each of these three variables loads rather lowly on the first dimension. Moreover, ANCOVA and MANOVA load similarly and somewhat highly (in the negative direction: see Table 3 and bottom-right portion of Figure 2) on the second dimension. This pattern appears to be contrasting the use of prediction methods and their accompanying weights and standard errors, with the use of group difference procedures of ANCOVA and MANOVA that do not traditionally report weights or standard errors.

Third, notice that a few of the multivariate methods (i.e., LR, MLM and Other) are positioned close to the middle of the graph, as are many of the inferential procedures (i.e., figures, fit indices, multiple models, exact *p*-values, CIs, mediation or moderation, longitudinal analysis, power analysis, and meta-analysis). This middle portion of the graph indicates low loadings for these variables, such that these methods and inferential procedures were not as salient in this two-dimensional CA solution. This finding probably reflects the fact that the set of European psychology journals sampled did not feature these as much in the articles we surveyed.

Contributions and Benefits

The current study offered contributions and benefits with a major investigation into the extent and nature of multivariate methods and inferential procedures used in a purposive sampling of eight major European psychology journals. Findings from the first question provide a view of recent practices in these eight European substantive psychology journals, with multivariate use reported in more than half of the articles with descending order of importance occurring for clinical, social, health, personality, organizational, developmental, educational, and cognitive articles, respectively. The number of multivariate articles across the eight journals varied widely with the highest use (i.e., in the Journal of Child Psychology and Psychiatry) showing more than eight times the amount of use in the least multivariate-based journal sampled here (i.e., the European Journal of Cognitive Psychology). Although causality cannot be inferred, for the specific set of journals surveyed it was notable that multivariate use was more strongly linked with higher-impact journals than with the lower impact journals. This issue could be explored further in a future study.

The second question results revealed that the kind of multivariate use in these journals involved predominatly prediction methods (i.e., multiple regression, structural equation modeling, and logistic regression), with some dimensional (i.e., factor or principal components analysis) and other kinds of multivariate analyses (e.g., cluster analysis, generalized estimating equations, genetic models, hazard analysis, multidimensional scaling); and less use of group difference methods (i.e., ANCOVA and MANOVA) and multilevel modeling. The latter finding is surprising as the development and study of multilevel modeling is evident among European methodological centers and institutions (e.g., University of Bristol in the UK; University of Oxford in the UK; Utrecht University, The Netherlands). It may be that whereas some methods such as multilevel modeling are highly studied by statistical researchers in Europe, they may be less apt to be adopted by substantive researchers. Alternatively, it may be that European researchers submit articles with rigorous methodology elsewhere (e.g., American journals).

The third question findings featured the kind of inferential procedures reported in these journals with strong attention to correlational or prediction effect sizes such as weights, and macro-level effect sizes (e.g., percentage of shared variance), as well as considerable use of traditional significance tests and *p*-values. On the one hand, such emphasis is at least partially consistent with current calls and recommendations for statistical reform (e.g., Fidler & Cumming, 2013; Thompson, 1996) including the use of effect sizes (e.g., Alhija & Levy, 2009; Cohen, 1992; Cumming, 2012; Grissom & Kim, 2012; Huberty, 2002; Kirk, 1996), among other procedures. On the other hand, the continued presence of dichotomous *p*-values and significance tests, although endorsed by some in the literature (e.g., Abelson, 1997; Chow, 1996; Robinson &

Levin, 2010), is somewhat surprising given encouragement to consider a wider range of inferential methods (e.g., Denis, 2003; Harlow, 2010; Harlow, Mulaik, & Steiger, 1997; Kline, 2004), and the stark discouragement of null hypothesis significance testing by others (e.g., Schmidt & Hunter, 1997; Thompson, 1996). Moreover, the relatively low use of confidence intervals (i.e., 17% of articles reported here), mediation and moderation (i.e., 16% reported here), and longitudinal data (i.e., 10% reported in these articles), as well as power- and meta-analyses (i.e., only 1% reported for these in this study) indicates that researchers are still not readily adopting recommended alternatives that could provide more informative scientific inferences (e.g., Cohen, 1988; Cumming, 2012; MacKinnon, 2008; Singer & Willet, 2003).

The fourth question brought together the type of multivariate methods used across the eight journals. When viewing the results, however, it is important to consider that researchers are more apt to select statistical analyses that are more well-known and for which they have sufficient background and training. The predominance of multiple regression use, particularly in clinical, social and health journals, and also in the other five European journals examined here, is probably due to wide knowledge and ease of learning about this method. This may also be true for structural equation modeling, which, although it is more complicated than multiple regression, is discussed in a large number of forums (e.g., books, journals, workshops, interest groups) and has reasonably accessible software associated with it (e.g., Amos, EQS).

Results for the fifth question relied on a correspondence analysis to explore two dimensions among the eight multivariate methods and the 14 inferential procedures examined in this study. The pattern of loadings for the first dimension revealed correspondence between often used structural equation modeling, significance tests, and dichotomous *p*-values, and between factor and principal components analysis and the percentage of variance explained. The second dimension demonstrated correspondence between highly used multiple regression and ES weights, with some standard errors use, in contrast with less use of ANCOVA and MANOVA.

Limitations

Despite the extent of the study and findings, several limitations are noted. First, results are not necessarily representative of all European journals as those from other journals may reveal a different pattern of multivariate inferential use. The current sample is more purposive than random, and journals from different years, impact factors, content areas, geographic areas, different schedules of publishing, and different considerations such as special issues or article-grouping journal practices could certainly affect results (many thanks to an anonymous reviewer for suggestions here). A second limitation is that many of the multivariate methods and inferential procedures had relatively low endorsement such that the correspondence analyses for the largely truncated variables assessing their use may not accurately capture findings due to the uneven distributions. Third, whether the nature of special issues of journals has any effect on the type of analyses used should be studied more closely. For example, in the current study several of the volumes in the set of journals examined addressed special issues. Fundamental Dimensions of Social Judgment was a special issue topic in which five of the articles from Issue 5 used multivariate analyses, of the total of 55 presented in the full Volume 38 of the December 2008 *European Journal of Social Psychology*. In another instance, the topic of Beyond Conscientiousness: A Personality Perspective on the Widening Sex Difference in School Performance was examined in Issue 3 of Volume 22 for the May 2008 European Journal of Psychology. In this case, a similarly reasonable number of seven of the special issue analyses were multivariate, of the full set of 43 multivariate analyses used in this volume. On the other hand, European Personality Reviews was the topic of a special Issue 5 in Volume 22 of this same journal, and none of these special issue articles used multivariate methods. Thus, the nature and extent of any special issues in journals needs further exploration in the future. Finally, although the current study was very time consuming to code and encompassed a broad and extensive survey of multivariate statistical inference procedures, information was not collected on several potentially relevant issues such as: whether the statistical methods used are appropriate given the specific research questions examined; the nature of the design (e.g., experimental, correlational, etc.); sample sizes of the studies; and size of the effects reported, to name a few.

Future Directions

Several suggestions are offered to build on current findings. For example, future research would be helpful to consider a wider and different range of journals, conducted in different content areas, countries, and years. It would also be useful to explore how closely the statistical methods used by researchers appeared to successfully address questions asked in each study, although this would be an extensive undertaking. Methodological researchers should be encouraged to publish understandable Teacher's Corner articles across a spectrum of substantive journals, and clearly written methodological application books, to provide guidelines on how to use rigorous multivariate methods in many areas of study.

Further, journal editors could call for a wider range of inferential procedure use in their journals (See, e.g., Fidler, Thomason, Cumming, Finch, & Leeman, 2004; La Greca, 2005; Thompson, 1996; Wilkinson, & The Task Force on Statistical Inference, 1999; Zedeck, 2003).

Finally, computer software distributers and developers should be encouraged to provide a wider range of options in their statistical routines that don't just focus mainly on significance tests and *p*-values. For example, although shared variance effect sizes are readily available in most multiple regression programs, they are not as likely to be easily obtainable in logistic regression and multilevel modeling programs. Similarly, confidence intervals are often provided for odds ratios in logistic regression, and root mean square error of approximation estimates in structural equation modeling programs; although more work could be done to provide accessible and informative indications of uncertainty for other parameter estimates and procedures.

Conclusions

In summary, multivariate methods and inferential procedures appear to be used in more than half of the 456 articles published in the eight European psychology journals sampled in the current study. There are differences in the nature and extent of multivariate methods used with half of the multivariate analyses involving multiple regression or structural equation modeling. Moreover, more than half of these analyses emphasized the use of micro-level ES weights, dichotomous *p*-values, figure diagrams, and significance tests. Future studies could examine whether there are similar practices in other journals, as well as how appropriate the analyses appeared to be for the research questions asked of the data. Lastly, further encouragement and illumination on how to use these encompassing and informative statistical inference tools could be beneficial in bringing about more insightful and revealing research findings that may lead to a more accurate and replicative body of psychological science.

References

Abelson, R. P. (1997). The surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science, 8*, 12-15. doi: 10.1111/j.1467-9280.1997.tb00536.x

Aberson, C. (2010). *Applied power analysis for the behavioral sciences.* New York: Routledge. ISBN: 978-1-84872-835-6

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317-332. doi: 10.1007/BF02294359

Alhija, F. N. -A. & Levy, A. (2009). Effect size reporting practices in published articles. *Educational and Psychological Measurement, 69*, 245-265. doi: 10.1177/0013164408315266

APA. (2010). *Publication manual of the American Psychological Association (6th Ed.*). Washington, DC: American Psychological Association. ISBN-10: 1-4338-0561-8

Balluerka, N., Gómez, J., & Hidalgo, D. (2005). The controversy over null hypothesis significance testing revisited. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1,* 55-70. doi: 10.1027/1614-1881.1.2.55

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley and S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499-525). San Diego: Academic Press. doi: 10.1016/B978-012691360-6/50018-5

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606. doi: 10.1037/0033-2909.88.3.588

Bodnar, O., Bodnar, T., & Gupta, A. K. (2010). Estimation and inference for dependence in multivariate data. *Journal of Multivariate Analysis, 101,* 869-881. doi: 10.1016/j.jmva.2009.11.005

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* New York: Wiley. ISBN: 978-0-470-05724-7

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd edition). New York: Springer. ISBN-10: 0-387-95364-7

Cafri, G., Kromrey, J. D., & Brannick, M. T. (2010). A meta-meta-analysis: Empirical review of statistical power, type I error rates, effect sizes, and model selection of meta-analyses published in psychology. *Multivariate Behavioral Research, 45*, 239-270. doi: 10.1080/00273171003680187

Chan, M. E., & Arvey, R. D. (2012). Meta-Analysis and the development of knowledge. *Perspectives on Psychological Science, 7*, 79-92. doi: 10.1177/1745691611429355

Chow, S. L. (1996). *Statistical significance: Rationale, validity and utility.* Beverly Hills, CA: Sage. ISBN: 0-920490-68-9

Cleveland, W. S. (1984). Graphical methods for data presentation: Full scale breaks, dot charts, and multibased logging. *The American Statistician, 38*, 270-280. doi: 10.1080/00031305.1984.10483224

Cleveland, W. S. (1993). *Visualizing data.* Murray Hill, NJ: AT&T Bell Laboratories. ISBN: 0-9634884-0-6

Cohen, J. (1962). The statistical power of abnormal-social psychological research. *Journal of Abnormal and Social Psychology, 65,* 143-153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^{nd} ed.). San Diego, CA: Academic Press. ISBN-10: 0805802835

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi: 10.1037/0033-2909.112.1.155

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 49*, 997-1003. doi: 10.1037/0003-066X.49.12.997

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cumming, G. (2008). Replication and *p* intervals: *p* values predict the future only vaguely, but confidence intervals do much better. *Perspectives on Psychological Science, 3*, 286-300. doi: 10.1111/j.1745-6924.2008.00079.x

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York: Routledge. ISBN: 978-0-415-87968-2

Denis, D. (2003). Alternatives to null hypothesis significance testing. *Theory & Science*, *4*, 1-21. Retrieved 9/8/11 from: (http://theoryandscience.icaap.org/content/vol4.1/02_denis.html).

Fidler, F. & Cumming, G. (2013). Effect size estimation and confidence intervals. In J. A. Schinka, W. F. Velicer, & I. B. Weiner (Eds.), *Handbook of psychology, Vol. 2: Research methods in psychology* (2nd ed.), (pp. 142-163). Hoboken, NJ: John Wiley & Sons Inc. ISBN: 978-0-470-89064-6

Fidler, F., Thomason, N., Cumming, G., Finch, S., & Leeman, J. (2004). Editors can lead researchers to confidence intervals, but can't make them think: Statistical reform lessons from medicine. *Psychological Science, 15,* 119-126. doi: 10.1111/j.0963-7214.2004.01502008.x

Friendly, M. (2000). *Visualizing categorical data.* Cary, NC: SAS Institute, Inc. doi: 10.1002/jhbs.20078

Glass, G. V. (1976). Primary, secondary and meta-analysis research. *Educational Researcher, 5*, 3-8. doi: 10.3102/0013189X005010003

Greenacre, M., & Hastie, T. (1987). The geometric interpretation of correspondence analysis. *Journal of the American Statistical Association, 82,* 437-447.

Grimm, L. G., & Yarnold, P. R. (Eds.) (1995). *Reading and understanding multivariate statistics.* Washington, DC: American Psychological Association. ISBN: 1-55798-273-2

Grimm, L. G., & Yarnold, P. R. (Eds.) (2000). *Reading and understanding more multivariate statistics.* Washington, DC: American Psychological Association. ISBN: 1-55798-698-3

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd edition). New York: Routledge. ISBN: 978-0-415-87769-5

Harlow, L. L. (2005). *The essence of multivariate thinking: Basic themes and methods.* Mahwah, NJ: Erlbaum. ISBN: 0-8058-3730-2

Harlow, L. L. (2010). On scientific research: Invited commentary on the role of statistical modeling and hypothesis testing. *Journal of Modern Applied Statistical Methods, 9,* 348-358.

Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum. ISBN: 0-8058-2634-3

Hess, M. R., Hogarty, K. Y., Ferron, J. M., & Kromrey, J. D. (2007). Interval estimates of multivariate effect sizes: Coverage and interval width estimates under variance heterogeneity and nonnormality. *Educational and Psychological Measurement*, *67*, 21-40. doi: 10.1177/0013164406288159

Huberty, C. J. (2002). A history of effect size indices. (Statistical Data Included). *Educational and Psychological Measurement*, *62*, 227-240. doi: 10.1177/0013164402062002002

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oakes, CA: Sage. ISBN-10: 141290479X

Kelley, K., & Preacher, K. J. (2012). On defining effect size. *Psychological Methods, 17,* April 30, 2012, No Pagination Specified. doi: 10.1037/a0028086.

Kirk, R. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56*, 746-759. doi: 10.1177/0013164496056005002

Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association. doi: 10.1037/10693-000

La Greca, A. M. (2005). Editorial. *Journal of Consulting and Clinical Psychology, 73,* 3-5. doi: 10.1037/0022-006X.73.1.3

Lai, K., & Kelley, K. (2011). Accuracy in parameter estimation for targeted effects in structural equation modeling: Sample size planning for narrow confidence intervals. *Psychological Methods, 16,* 127-148. doi: 10.1037/a0021764

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* New York: Erlbaum. ISBN: 978-0-8058-6429-8

Maxwell, A. E. (1961). Canonical variate analysis when the variables are dichotomous. *Educational and Psychological Measurement, 21,* 259-271.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Erlbaum. ISBN: 0-8058-3718-3

McGrath, R. E. (2011). *Quantitative models in psychology.* Washington, DC: APA. doi: 10.1037/12316-000

Mendoza, J. L., & Stafford, K. L. (2001). Confidence intervals, power calculation, and sample size estimation for the squared multiple correlation coefficient under the fixed and random regression models: A computer program and useful standard tables. *Educational and Psychological Measurement, 61,* 650-667. doi: 10.1177/00131640121971419

Mulaik, S. A., Raju, N. S., & Harshman, R. A. (1997). There is a time and place for significance testing. In L. L. Harlow, S. A. Mulaik, and J. H. Steiger, *What if there were no significance testing?* (pp. 65-115). Mahwah, NJ: Erlbaum. ISBN: 0-8058-2634-3

Nam, I. –S., Mengersen, K., & Garthwaite, P. (2003). Multivariate meta-analysis. *Statistics in Medicine, 22,* 2309-2333. doi: 10.1002/sim.1410

Pearson, K. (1885). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 186*, 343–414. Retrieved April 6, 2012 from: http://rsta.royalsocietypublishing.org/content/186/343.full.pdf+html?sid=11e2cf1a-b1b5-4c6a-a5fb-71f514c3b3a9

Peng, C.-Y. J., Chen, L.-T., Chiang, H.-M., & Chiang, Y.-C. (2013). The impact of APA and AERA guidelines on effect size reporting. *Educational Psychology Review*, Mar 12, first posting: no pagination specified. doi:10.1007/s10648-013-9218-2

Raudenbush, S. W., Becker, B. J., & Kalaian, H. (1988). Modeling multivariate effect size. *Psychological Bulletin, 103,* 111-120. doi: 10.1037/0033-2909.103.1.111

Robinson, D. H., & Levin, J. R. (2010). The not-so-quiet revolution: Cautionary comments on the rejection of hypothesis testing in favor of a 'causal' modeling alternative. *Journal of Modern Applied Statistical Methods, 9*, 332-339.

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, *65*, 1-12. doi: 10.1037/a0018326

Rossi, J. S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology, 58,* 646-656.

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 37-64). Mahwah, NJ: Erlbaum. ISBN: 0-8058-3730-2

Singer, J. D., & Willett, J. B. (2003). *Applied l*ongitudinal *d*ata *a*nalysis: *Modeling change and e*vent *o*ccurrence. *New York: Oxford University Press.* ISBN 0-19-515296-4

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models.* Boca Raton, FL: Chapman & Hall. ISBN 9781584880004

Steiger, J. H., & Fouladi, R. T. (1992). R^2: A computer program for interval estimation, power calculation, sample size estimation, and hypothesis testing in multiple regression. *Behavior Research Methods, Instruments, and Computers, 24*, 581-582. doi: 10.3758/BF03203611

Steiger, J. H., & Lind, J. C. (1980, June). *Statistically based tests for the number of common factors.* Paper presented at the Psychometric Society annual meeting, Iowa City, IA.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Routledge. ISBN: 978-0-8058-5903-4

Steyn Jr., H. S., & Ellis, S. M. (2009). Estimating an effect size in one-way multivariate analysis of variance (MANOVA). *Multivariate Behavioral Research, 44*, 106-129. doi: 10.1080/00273170802620238

Sun, S., Pan, W., & Wang, L. L. (2011). Rethinking observed power: Concept, practice, and implications. *Methodology*, *7,* 81-87. doi: 10.1027/1614-2241/a000025

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate* statistics (6^{th} ed.). Boston: Pearson. ISBN-10: 0205459382

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, *25*, 26-30. doi: 10.3102/0013189X025002026

Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, *44*, 423-432. doi: 10.1002/pits.20234

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley. ISBN-10: 0201076160

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American P5sychologist*, *54,* 594–604.

Yazici, A. C., Öğüş, E., Ankarali, H., & Gürbüz, F. (2010). An application of nonlinear canonical correlation analysis on medical data. *Turkish Journal of Medical Science, 40,* 503-510. doi:10.3906/sag-0803-47

Van de Geer, J. P. (1987). Algebra and geometry of OVERALS. *Research Reports 87-13,* 1-47. University of Leiden: Department of Data Theory. Available from: URL: http://www.datatheory.nl/pdfs/87/87_13.pdf

Zedeck, S. (Ed) (2003). Applied psychology: Editorial. *Journal of Applied Psychology, 88,* 3-5. doi: 10.1037/0021-9010.88.1.3

Table 1
Selected Content Journals, Impact Factors, and Percentages of 2008 Articles Using Multivariate Methods.

Content Areas	Research Journals	[1]Impact Factor	# of 2008 Multivariate Articles	Total # of 2008 Research Articles	% of Multivariate Articles per Journal
Clinical	(European) Journal of Child Psychology and Psychiatry	4.36	83	143	58%
Personality	European Journal of Personality	2.10	25	30	83%
Health	(European Journal of) Psychology and Health	1.59	39	60	65%
Social	European Journal of Social Psychology	1.59	55	89	62%
Applied	European Journal of Work & Organizational Psychology	1.49	20	24	83%
Cognitive	European Journal of Cognitive Psychology	1.24	11	50	22%
Developmental	European Journal of Developmental Psychology	1.03	17	31	55%
Educational	European Journal of Psychology of Education	0.55	9	29	31%
	Average Impact Factor:	1.74			
	Totals of Articles:		259	456	
	Mean Percentage of Articles:				57%

[1]**Note:** Impact factors obtained from 2010 on-line Web of Science report that assesses the impact of articles published in the 2008 articles examined in this study. The correlation between impact factor and the number of 2008 multivariate articles reported is .86.

Table 2
Distributional Patterns across 8 Journals, 8 Multivariate Methods and 14 Inferential Procedures

1. 8 European Journals	%		**2. 8 Multivariate Methods**	%		**3. 14 Inferential Procedures**	%
Clinical	29.3		Multiple regression	33.7		Weights	63.8
Social	19.5		Structural models	17.0		Dichotomous *p*	56.2
Health	16.7		Logistic regression	9.8		Figures	55.3
Personality	11.8		FA-PCA	9.6		Significance tests	51.5
Organizational	8.2		ANCOVA	8.5		% Variance	45.8
Developmental	6.8		MANOVA	8.2		Fit indices	44.9
Educational	4.1		Other	7.7		Multiple models	43.6
Cognitive	3.6		MLM	5.5		Exact *p*	25.5
						Standard errors	23.3
						Confidence intervals	17.0
						Mediation-Moderation	15.9
						Longitudinal	9.6
						Meta-analysis	1.4
						Power analysis	0.8

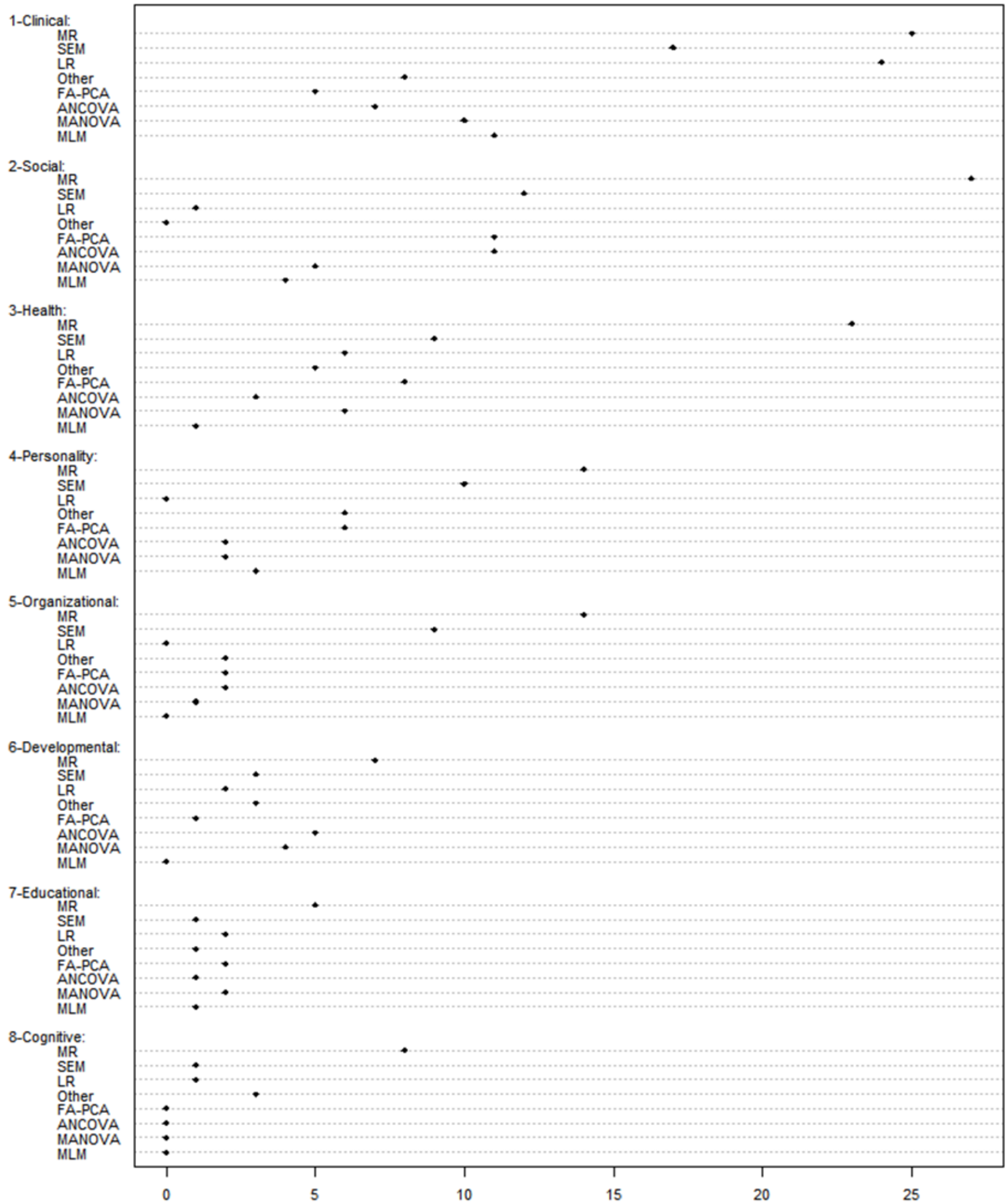
EUROPEAN JOURNALS Note: Clinical = (European) Journal of Child Psychology and Psychiatry, Social = European Journal of Social Psychology, Health = (European Journal of) Psychology and Health, Personality = European Journal of Personality, Organizational = European Journal of Work and Organizational Psychology, Developmental = European Journal of Developmental Psychology, Educational = European Journal of Psychology of Education, Cognitive = European Journal of Cognitive Psychology. MULTIVARIATE METHODS Note: Other = cluster analysis, generalized estimating equations, genetic models, hazard analysis, multidimensional scaling, survival analysis, time series, etc. INFERENTIAL PROCEDURES Note: Percentages add to more than 100% as most analyses used several inferential procedures.

Table 3
Component Loadings for the 2 Dimensions from the Correspondence Analysis

Set	Variables	Dimension 1	Dimension 2
1	Ancova12	.240	**-.410**
	Manova12	.214	**-.413**
	LR12	.051	.058
	MR12	-.097	**.698**
	MLM12	-.046	.212
	FAPCA12	**-.797**	-.271
	SEM12	**.355**	-.024
	OthStat12	.204	-.180
2	Wts12	-.109	**.815**
	pLT12	**.375**	.254
	FigData12	.123	-.087
	X2orF12	**.571**	-.196
	PerVar12	**-.341**	.061
	AnyFit12	.153	-.129
	MM12	.293	.221
	pEQ12	.273	.046
	SE12	.079	**.347**
	CI12	.191	-.005
	MedMod12	.123	.287
	Longit12	.163	.024
	Meta12	.095	-.062
	Powr12	.017	.025

Note: MR = multiple regression, SEM = structural equation modeling, LR = logistic regression, FA-PCA = factor analysis or principal components analysis, Other = cluster analysis, generalized estimating equations, genetic models, hazard analysis, multidimensional scaling, survival analysis, time series, etc.; Ancova = analysis of covariance, Manova = multivariate analysis of variance, MLM = multilevel modeling. Wts = standardized or unstandardized weights, loadings, odds ratios, etc; pLT = use of $p < .05$, etc.; FigData = figures with data; X2orF = significance tests; PerVar = percentage of shared variance; AnyFit = supplemental fit index; MM = multiple models; pEQ = use of $p = .03$, etc.; SE = standard errors; CI = confidence intervals; MedMod = mediation or moderation analysis; Longit = longitudinal analysis; Meta = meta-analysis; and Powr = statistical power. The end-labels “12” indicate values of 1 and 2, instead of 0 and 1, for CA. Bolded values indicate most salient component loadings on the two CA dimensions.

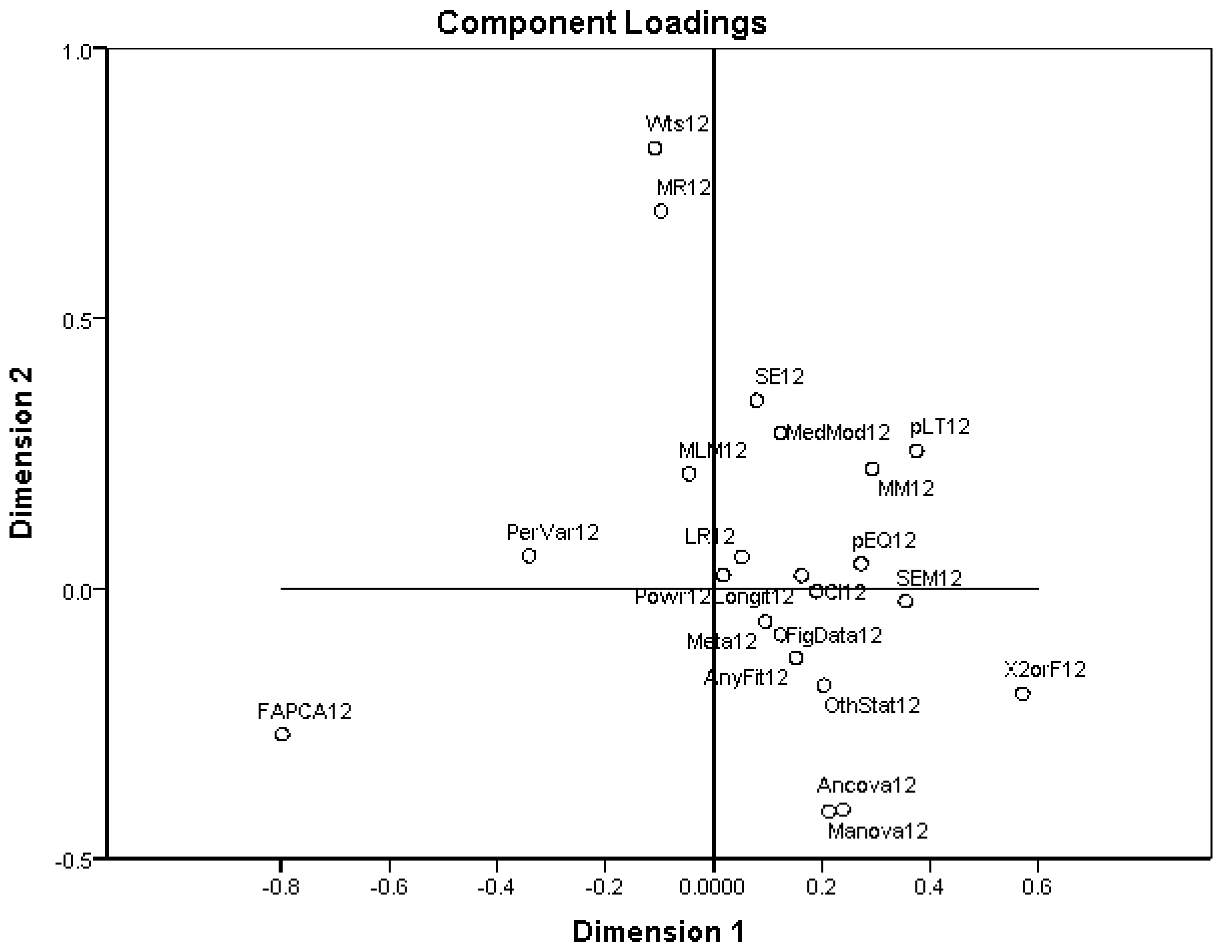
Figure 1
Dot plot of Frequency of Analyses Conducted with Each of 8 Kinds of Multivariate Methods, Over the 8 Different European Psychology Journals (Listed by Content Focus)



Note: MR = multiple regression; SEM = structural equation modeling; LR = logistic regression; Other = cluster analysis, generalized estimating equations, genetic models, hazard analysis, multidimensional scaling, survival analysis, time series, etc.; FA-PCA = factor analysis or principal components analysis; ANCOVA = analysis of covariance; MANOVA = multivariate analysis of variance; MLM = multilevel modeling. Clinical = (European) Journal of Child

Psychology and Psychiatry, Social = European Journal of Social Psychology, Health = (European Journal of) Psychology and Health, Personality = European Journal of Personality, Organizational = European Journal of Work and Organizational Psychology, Developmental = European Journal of Developmental Psychology, Educational = European Journal of Psychology of Education, Cognitive = European Journal of Cognitive Psychology.

Figure 2
Graph of Component Loadings of 8 Methods and 14 Inferential Procedures on 2 CA Dimensions



Note: MR = multiple regression, SEM = structural equation modeling, LR = logistic regression, FA-PCA = factor analysis or principal components analysis, Other = cluster analysis, generalized estimating equations, genetic models, hazard analysis, multidimensional scaling, survival analysis, time series, etc.; Ancova = analysis of covariance, Manova = multivariate analysis of variance, MLM = multilevel modeling. Wts = standardized or unstandardized weights, loadings, odds ratios, etc; pLT = use of $p < .05$, etc.; FigData = figures with data; X2orF = significance tests; PerVar = percentage of shared variance; AnyFit = supplemental fit index; MM = multiple models; pEQ = use of $p = .03$, etc.; SE = standard errors; CI = confidence intervals; MedMod = mediation or moderation analysis; Longit = longitudinal analysis; Meta = meta-analysis; and Powr = statistical power. End-labels “12” indicate values of 1 and 2, instead of 0 and 1, for CA.