Educational and Psychological Measurement
Volume 66 Number 3
June 2006 393-416

10.1177/0013164405282485
http://epm.sagepub.com
hosted at
http://online.sagepub.com

Use of Exploratory Factor Analysis in Published Research

Common Errors and Some Comment on Improved Practice

Robin K. Henson
University of North Texas
J. Kyle Roberts
Baylor College of Medicine

Given the proliferation of factor analysis applications in the literature, the present article examines the use of factor analysis in current published research across four psychological journals. Notwithstanding ease of analysis due to computers, the appropriate use of factor analysis requires a series of thoughtful researcher judgments. These judgments directly affect results and interpretations. The authors examine across studies (a) the decisions made while conducting exploratory factor analyses ($N = 60$) and (b) the information reported from the analyses. In doing so, they present a review of the current status of factor analytic practice, including comment on common errors in use and reporting. Recommendations are proffered for future practice as regards analytic decisions and reporting in empirical research.

Keywords: *exploratory factor analysis (EFA); factor retention; meta-analysis; confirmatory factor analysis (CFA)*

Researchers commonly attempt to explain the most with the least. For example, because all classical parametric analyses are part of a broader general linear model, these analyses are all correlational, yield r^2-type effect sizes, and maximize the shared variance between variables (e.g., regression) or between sets of variables (e.g., canonical correlation; Bagozzi, Fornell, & Larcker, 1981; Cohen, 1968; Henson, 2000; Knapp, 1978; Thompson, 1991). In fact, classical parametric analyses (e.g., ANOVA, *r*, MANOVA, descriptive discriminant analysis) can all be performed with canonical correlation analysis and, thus, are special cases of canonical analysis (Fan, 1996, 1997; Thompson, 2000a).

Because implicit within canonical correlation analysis itself is a principal components analysis (PCA; Thompson, 1984, pp. 11-14), all classical parametric analyses

Authors' Note: Please address correspondence to Robin K. Henson, Department of Technology and Cognition, University of North Texas, P.O. Box 311337, Denton, TX 76203-1337.

also invoke PCAs (albeit in different ways). This truism suggests the importance of factor analysis within statistics.

In the interest of parsimony, researchers often strive to explain the most shared variance of measured variables using the fewest possible latent or synthetic variables. Such parsimonious solutions are generally considered to have greater external validity and, as such, are more likely to replicate. Thus, Kerlinger (1979) argued that factor analysis is "one of the most powerful methods yet for reducing variable complexity to greater simplicity" (p. 180).

Factor analysis is often used to explain a larger set of j measured variables with a smaller set of k latent constructs. It is hoped, generally, that the k constructs will explain a good portion of the variance in the original $j \times j$ matrix of associations (e.g., correlation matrix) so that the constructs, or factors, can then be used to represent the observed variables. These constructs can be used as variables in subsequent analyses and "can be seen as actually causing the observed scores on the measured variables" (Thompson & Daniel, 1996, p. 202). In short, "factor analysis is intimately involved with questions of validity . . . [and] is at the heart of the measurement of psychological constructs" (Nunnally, 1978, pp. 112-113).

Historically, the theoretical framework for factor analysis is credited to Pearson (1901) and Spearman (1904), but practical application of the method is a modern phenomenon. As Kieffer (1999) noted,

> Spearman, through his work on personality theory, provided the conceptual and theoretical rationale for both exploratory and confirmatory factor analysis. Despite the fact that the conceptual bases for these methods have been available for many decades, it was not until the wide-spread availability of both the computer and modern statistical software that these analytic techniques were employed with any regularity. (p. 75)

Thanks to the advent of technology, factor analysis is now frequently employed in both measurement and substantive research.

Purpose

Given the proliferation of factor analysis applications in the literature, the purpose of the present article is to examine the use of factor analysis in current published research. Notwithstanding ease of analysis due to computers, the appropriate use of factor analysis requires a series of thoughtful researcher judgments. These judgments directly affect results and interpretations.

Specifically, we examine across studies (a) the decisions made while conducting exploratory factor analyses (EFAs) and (b) the information reported from the analyses. In doing so, we present here a review of the current status of factor analytic practice, including comment on common errors in use and reporting. We then make recommendations for future practice as regards analytic decisions and reporting in empirical research.

EFA

Modern conceptualizations of factor analysis include both exploratory and confirmatory methods, as well as hybrids invoking exploratory factor extraction followed by confirmatory rotation (Thompson, 1992) or confirmatory maximum likelihood factor analysis (Jöreskog, 1969). EFA is used to "identify the factor structure or model for a set of variables" (Bandalos, 1996, p. 389). As its name implies, EFA is an exploratory method used to *generate* theory; researchers use EFA to search for the smaller set of *k* latent factors to represent the larger set of *j* variables. As Pedhazur and Schmelkin (1991) noted, "Of the various approaches to studying the internal structure of a set of variables or indicators, probably the most useful is some variant of factor analysis" (p. 66).

On the other hand, confirmatory factor analysis (CFA) is generally used to *test* theory when the analyst has sufficiently strong rationale regarding what factors should be in the data and what variables should define each factor. A fundamental and critically important difference between EFA and CFA is that results of an EFA are a sole function of the "mechanics and mathematics of the method" (Kieffer, 1999, p. 77). Although the researcher may have some conceptualization of what factors may be present in the data, such as when items are developed to measure expected constructs, EFA generally does not consider strong a priori theory (Daniel, 1989). CFA, on the other hand, is typically driven by theoretical expectations regarding the structure of the data.

As Gorsuch (1983) noted, "Whereas the former [EFA] simply finds those factors that best reproduce the variables under the maximum likelihood conditions, the latter [CFA] tests specific hypothesis regarding the nature of the factors" (p. 129). The reader is referred to Gorsuch, Stevens (1996), Tabachnick and Fidell (1996), and Thompson (2004) for extensive treatments of these approaches. The present article is concerned with EFA.

Purposes of Factor Analysis

Because the latent constructs, or factors, are thought to cause and summarize responses to observed variables, theory development and score validity evaluation are both closely related to factor analysis. As Hendrick and Hendrick (1986) emphasized, "Theory building and construct measurement are joint bootstrap operations" (p. 393). Factor analysis at once both tests measurement integrity and guides further theory refinement.

As noted by Kieffer (1999), "the utilization of factor analytic techniques in the social sciences has been indelibly intertwined with [both] developing theories and evaluating the construct validity of measures [i.e., scores]" (p. 75). Regarding construct validity, Gorsuch (1983) noted,

> Research proceeds by utilizing operational referents for the constructs of a theory to test if the constructs interrelate as the theory states. . . . A prime use of factor analysis has been in the development of both the operational constructs for an area and the operational representatives for the theoretical constructs. (p. 350)

Factor analysis can be used to determine what theoretical constructs underlie a given data set and the extent to which these constructs represent the original variables. Of course, the meaningfulness of latent factors is ultimately dependent on researcher definition. As Mulaik (1987) suggested, "It is we who create meanings for things in deciding how they are to be used. Thus we should see the folly of supposing that EFA will teach us what intelligence is, or what personality is" (p. 301). However, Thompson and Daniel (1996) noted that

> analytic results can inform the definitions we wish to create, even though we remain responsible for our elaborations and may even wish to retain the definitions that have not yet been empirically supported or that limited empirical evidence may even contradict. (p. 202)

(Thoughtful) Researcher Judgment in EFA

Despite its utility in both measurement and substantive research contexts, factor analysis has been criticized. Cronkhite and Liska (1980) observed,

> Apparently, it is so easy to find semantic scales which seem relevant, . . . so easy to name or describe potential/hypothetical sources, so easy to capture college students to use the scales to rate the sources, so easy to submit those ratings to factor analysis, so much fun to name the factors when one's research assistant returns with the computer printout, and so rewarding to have a guaranteed publication with no fear of nonsignificant results that researchers, once exposed to the pleasures of the factor analytic approach, rapidly become addicted to it. (p. 102)

Much of the criticism centers on the inherent subjectivity of the decisions necessary to conduct an EFA. Tabachnick and Fidell (1996) stated that "one of the problems with [PCA] and [factor analysis] is that there is no criterion variable against which to test the solution" (p. 636). Interpretation of results largely hinges on (it is hoped) reflective researcher judgment. Tabachnick and Fidell also noted that after factor extraction,

> there is an infinite number of rotations available, all accounting for the same amount of variance in the original data, but with factors defined slightly differently. The final choice among alternatives depends on the researcher's assessment of its interpretability and scientific utility. In the presence of an infinite number of mathematically identical solutions, researchers are bound to differ regarding which is best. Because the differences cannot be resolved by appeal to objective criteria, arguments over the best solution sometimes become vociferous. (p. 636)

Because EFA

> can be conceptualized as a series of steps which require that certain decisions be addressed at each individual stage . . . there are many different ways in which to conduct an EFA, and each different approach may render distinct results when certain conditions are satisfied. (Kieffer, 1999, pp. 76-77)

Therefore, appropriate use of EFA necessitates thoughtful and informed researcher decision making.

EFA Decisions

A complete review of the steps and possible decisions necessary to conduct an EFA is beyond the scope of this article. However, a brief review is given here to place the current study in context. Comprehensive treatments are provided by Gorsuch (1983) and Thompson (2004). Hetzel (1996) and Kieffer (1999) presented briefer user-friendly primers on factor analysis.

Matrix of Associations

Because all classical statistical analyses are fundamentally correlational (cf. Cohen, 1968; Knapp, 1978), all analyses focus on a matrix of associations that describes the relationships between the variables in question. To conduct an EFA, the researcher must decide which matrix of associations (e.g., correlation, variance/covariance) to analyze. Most statistical packages use the correlation matrix as the default option in EFA. Subsequently, researchers tend to use the correlation matrix.

Method of Factor Extraction

There are multiple ways to extract factors. PCA and principal axis factoring (PAF) tend to be the most common. Factor extraction attempts to remove variance common to sets of variables from the original matrix of association. After the first factor (or common variance for a set of variables) has been extracted, a residual matrix remains. A second factor, which is orthogonal to the first, is then extracted from the residual matrix to explain as much of the remaining variance among the variables as possible. The process continues until noteworthy variance can no longer be explained by factors.

The application of PCA against PAF has been hotly debated. As Thompson and Daniel (1996) noted,

> Analysts differ quite heatedly over the utility of principal components as compared to common or principal factor analysis [i.e., PAF]. . . . The differences between the two approaches involve the entries used on the diagonal of the matrix of associations that is

> analyzed. When a correlation matrix is analyzed, principal components analysis uses ones on the diagonal whereas common factor analysis uses estimates of reliability, usually estimated through an iterative process. (p. 201)

Gorsuch (1983) suggested that the researcher consider carefully which method to use because differences can be meaningful.

Indeed, many factor analysts would not consider PCA a factor analysis at all. Instead, PCA is intended to simply summarize many variables into fewer components, and the latent constructs (i.e., factors) are not the focus of the analysis. On the other hand, PAF explicitly focuses on the common variance among the items and, therefore, focuses on the latent factor. Fabrigar, Wegener, MacCallum, and Strahan (1999) provided a more complete discussion of this issue. This distinction may be important in some contexts; however, for ease of communication, we have elected to refer to both PCA and PAF as factor analyses. In our subsequent review of the literature, we do indicate which specific extraction method was employed.

Thompson (1992) argued that the practical difference between the methods is often negligible in terms of interpretation. Differences in results will decrease as (a) the measured variables have greater score reliability or (b) the number of variables measured increases. Regarding (a), the higher the score reliability for a variable, the closer the PAF entry on the diagonal is to one, which is what is used by PCA.

Regarding (b), as the number of variables increases, so does the total number of entries on the matrix of association. The influence of the diagonal entries then has less influence on the solution because the proportion of entries on the diagonal decreases exponentially as more variables are measured (cf. Snook & Gorsuch, 1989). For example, with 10 measured variables, there are 10 diagonal elements in the matrix of association out of 55 nonredundant entries (i.e., 18.2%); with 30 measured variables, there are 30 diagonal elements out of 435 nonredundant entries (i.e., 6.9%); and with 50 measured variables, there are 50 diagonal elements out of 1,225 total entries (i.e., 4.1%).

Factor Retention Rules

When variables are factored (for a discussion of factoring people, see Campbell, 1996; Thompson, 2000b), the total number of possible factors equals the number of variables factored (assuming all of the variance in the original variables is not reproduced). However, because many of these factors may not contribute substantially to the overall solution or be interpretable, some factors are not useful to retain in the analysis and generally represent noise or error. Given that the goal of EFA is to retain the fewest possible factors while explaining the most variance of the observed variables, it is critical that the researcher extract the correct number of factors because this decision will affect results directly.

Many rules can be used to determine the number of factors to retain (cf. Zwick & Velicer, 1986), including the eigenvalue > 1 rule (EV > 1; Kaiser, 1960), scree test (Cattell, 1966), minimum average partial correlation (Velicer, 1976), Bartlett's chi-square test (Bartlett, 1950, 1951), and parallel analysis (Horn, 1965; Turner, 1998). Thompson and Daniel (1996) and Zwick and Velicer (1986) elaborated these approaches. The most frequently used method is the EV > 1 rule. As Thompson and Daniel noted, "This extraction rule is the default option in most statistics packages and therefore may be the most widely used decision rule, also by default" (p. 200).

Importantly, these rules do not necessarily lead to the same decision regarding the number of factors to retain. For example, in a Monte Carlo evaluation, Zwick and Velicer (1986) found that the EV > 1 rule almost always severely overestimated the number of factors to retain. Their findings were consistent with those of Cattell and Jaspers (1967), Linn (1968), Yeomans and Golder (1982), and Zwick and Velicer (1982), but they were contrary to those of Humphreys (1964) and Mote (1970), who noted that the EV > 1 rule may underestimate the number of factors.

Bartlett's chi-square test was very inconsistent. Because EFA studies typically involve large samples, this statistical significance test may have little utility as it is heavily influenced by sample size.

Despite its subjective nature in interpretation, the scree test was much more accurate but also tended to overextract factors. Importantly, parallel analysis was the most accurate procedure, followed closely by the minimum average partial method. Unfortunately, these methods are seldom employed in published research. As an additional option, Thompson (1988) suggested using a bootstrap method to determine the number of factors and provided a program to automate the process.

Because the factor retention decision directly affects the EFA results obtained, researchers are advised to use both multiple criteria and reasoned reflection. Researchers should also explicitly inform readers about the strategies used in making factor retention decisions.

Factor Rotation and Coefficient Interpretation

Rotation strategies are numerous and can be classified into two broad categories: orthogonal and oblique. Almost all researchers rotate their EFA results to facilitate interpretation of their factors. Discussion of the various rotation strategies is dealt with elsewhere (cf. Gorsuch, 1983; Kieffer, 1999; Stevens, 1996; Thompson, 2004) and will not be addressed here. However, one point will be made regarding the coefficients used when interpreting EFA results.

In EFA, the contribution of a variable to a given factor is indicated by both factor pattern coefficients (sometimes ambiguously called "loadings") and factor structure coefficients (also sometimes ambiguously called "loadings"). Thompson and Daniel (1996) noted that structure coefficients, or correlations between observed and latent variables, "are usually essential to interpretation" (p. 199). Their sentiment applies not only to factor analysis but also to other general linear model analyses (cf. Henson, 2002; Thompson & Borrello, 1985).

In factor analysis, the factor structure matrix gives the correlations between all observed variables and all extracted (latent) factors. When factors are orthogonally rotated, they remain uncorrelated, and the factor structure matrix will exactly match the factor pattern matrix. Mathematically, the structure matrix is obtained by multiplying the factor pattern matrix ($\boldsymbol{P}_{V\times F}$) by the factor correlation matrix ($\boldsymbol{R}_{F\times F}$), which isan identity matrix (i.e., ones on diagonal, zeros off diagonal) after orthogonal rotation. The resulting structure matrix ($\boldsymbol{S}_{V\times F}$) will match the original factor pattern matrix (cf. Gorsuch, 1983, p. 52) whenever the factor correlation matrix is an identity matrix. In such cases, the pattern matrix should be called the "factor pattern/structure matrix" to facilitate clarity. Because *loading* is used ambiguously in the literature, use of this term is proscribed by some editorial policies (Thompson & Daniel, 1996).

When an oblique rotation is used, the factors are allowed to correlate with each other. In such cases, the factor correlation matrix will not be an identity matrix, and the structure matrix will not equal the pattern matrix. Appropriate interpretation, then, must invoke both the factor pattern and factor structure matrices. Because all analyses are correlational and belong to the general linear model, the problem of only interpreting the factor pattern coefficients is analogous to only interpreting beta weights in regression when the predictors are correlated. As illustrated by Courville and Thompson (2001), Henson (2002), and Thompson and Borrello (1985), consideration of structure coefficients is critical in such cases.

EFA Reporting Practices

Replication is a foundational principle of science (Henson & Smith, 2000; Thompson, 1999). Findings in a single study seldom "prove" anything, but confidence in results increases when independent researchers externally evaluate the validity of previously reported research. Regarding factor analysis, it is very important that researchers be able to independently evaluate the results obtained in an EFA study. This can, and should, occur on two levels. Given the myriad subjective decisions necessary in EFA, independent researchers should be able to evaluate the analytic choices of authors in the reported study. Second, independent researchers should be able to accurately replicate the study on new data, or even employ a CFA. (Of course, a secondary CFA should be conducted on a new sample. It is not terribly informative, and can be potentially misleading, to follow an EFA with a CFA on the same data set.)

Unfortunately, such practices are not possible for most applications of EFA. As noted by Tinsley and Tinsley (1987), most applied uses of factor analysis do not provide sufficient information to allow others to make independent interpretations. Too often, authors report only the final results of their factor analysis, thereby eliminating the possibility of external evaluation of EFA decisions (cf. Comrey, 1978). In addition, some authors do not report sufficient information to even allow independent interpretation of the final results, such as only giving part of the factor pattern matrix or excluding the factor structure matrix for oblique solutions (Thompson & Daniel, 1996; Tinsley & Tinsley, 1987). Many authors have called for more detailed factor analytic information in published research (cf. Comrey, 1978; Gorsuch, 1983; Henson, Capraro, & Capraro, 2004; Kline, 1994; Thompson & Daniel, 1996; Tinsley & Tinsley, 1987; Weiss, 1971).

In a review of 13 factor analysis articles in the *Journal of Counseling Psychology*, Hetzel (1996) found that much of this information was not reported by authors. Fabrigar et al. (1999) also provided a review of two journals and noted frequent poor decisions and missing information. Of course, precious journal space may limit information given, but critical decisions should nevertheless be explicitly addressed (e.g., the rule(s) used to determine the number of factors to retain), and complete information should be made available to interested persons (cf. Tinsley & Tinsley, 1987).

As noted, the purpose of the present article was to examine the EFA decisions and reporting practices in published EFA research. Although Fabrigar et al. (1999) and Hetzel (1996) characterized basic patterns of reporting in the counseling and psychology literature, the present review (a) broadened the literature studied to include both measurement and substantive articles from four different journals and (b) considerably expanded the reporting practices and decisions examined.

Method

Journal and Article Selection

Journals frequently employing factor analytic studies were identified from a search of the ERIC and PsycLIT databases using the keywords *factor analysis*. Although many journals publish articles using EFA, the following four journals were selected for investigation because of their greater reporting frequency as regards EFA applications: *Educational and Psychological Measurement*, *Journal of Educational Psychology*, *Personality and Individual Differences*, and *Psychological Assessment*. These journals also reflect both substantive and measurement applications of EFA.

Fifteen uses of EFA were examined from each of the journals, resulting in 60 total EFAs studied. Articles were selected if they employed EFA; articles using only CFA were not examined. In addition, if 1 article included more than 1 EFA, all EFAs were coded if they were substantively different in terms of the information reported. We began examining articles from the end of 1999 (except for *Personality and Individual Differences*, for which only articles from Vol. 26, June 1999, and earlier were available) and worked backwards until 15 applications of EFA from each journal were identified. A total of 432 articles were examined. Forty-nine articles were identified that used 1 or more EFAs, giving a total EFA sample of 60. These EFAs were coded on multiple criteria to assess the information reported and the analytic decisions made by authors.

Results and Discussion

Table 1 presents descriptive results for six global EFA variables. The sample size distribution was quite variable and positively skewed (coefficient of skewness = 3.07). The median sample size (267) would be classified as somewhere between fair and good according to Comrey and Lee (1992), who portrayed as a guide size of 50 as very

Table 1
General Descriptive Results of Exploratory Factor Analysis Reporting Practices

Variable	*n*	Median	*M*	*SD*	Minimum	Maximum
Sample size	59	267.00	436.08	540.74	42	3,113
Ratio of no. of participants to no. of variables factored	59	11.00[a]	26.86	52.79	3.25	348.40
No. of variables factored	60	20.00	23.73	16.70	5	110
No. of factors extracted	60	3.00	3.48	1.46	1	7
Cutoff used to determine what coefficients were meaningfully weighted on a factor	37	0.40	0.40	0.07	0.30	0.50
Total variance explained by extracted factors	43	51.70%	52.03%	14.48%	16.70%	87.50%

Note: *n* = number of articles reporting the relevant information.
a. Indicates that there were 11 participants per one variable factored.

poor, 100 as poor, 200 as fair, 300 as good, 500 as very good, and 1,000 as excellent. It is not uncommon to find such rules of thumb in the factor analytic literature; it is less common, though, to find consistency in recommendations. For example, Tabachnick and Fidell (1996) noted that "as a general rule of thumb, it is comforting to have at least 300 cases for factor analysis" (p. 640).

Stevens (1996) suggested that the number of participants per variable is a more appropriate way to determine sample size (ranging from 5 to 20 participants per variable). Fewer participants are needed when component saturation is high. In the current sample, the median ratio of number of participants to variables was 11:1 (coefficient of skewness = 4.67), suggesting that most sample sizes were marginal to sufficient, depending on component saturation. However, 7 EFAs (11.86%) had ratios less than Stevens's minimum of 5:1. One study failed to report sample size.

In general, however, sample size rules of thumb fail to take into account many of the complex dynamics of a factor analysis. MacCallum, Widaman, Zhang, and Hong (1999) demonstrated that such rules of thumb can at times be misleading. Instead, the adequacy of a sample size depends in large part to the features of the obtained data, which means that definitive a priori decisions about sample size can be difficult. MacCallum et al. illustrated that when communalities are high (greater than .60) and each factor is defined by several items (typically four or more), sample sizes can actually be relatively small. When communalities are not high, however, larger sample sizes are needed. Because one cannot know for sure how strong the communalities will be until the data are analyzed, perhaps the best rule of thumb to follow is to get the largest possible sample for a factor analysis.

The extracted factors explained, on average, 52.03% of the total variance in the original variables. This amount is drastically less than the "75% or more" recommended by Stevens (1996, p. 364). It is also inconsistent with Gorsuch's (1983) claim that "usually, investigators compute the cumulative percentage of variance extracted after each factor is removed from the matrix [of association] and then stop the factoring process when 75, 80 or 85% of the variance is accounted for" (p. 165). Only the most effective EFAs in the current study met these criteria for variance accounted for. It is unclear whether the modest explained variance was owing to researchers' failing to extract meaningful factors in their data or that their instruments failed to yield data with clear internal structure that can be represented by latent constructs. The current finding also calls into question whether 75% or more variance explained is a reasonable expectation in applied psychological research.

As may be expected, the proportion of total variance explained tended to decrease as the total number of items factored increased ($r = -.113$, $n = 43$). This pattern was also consistent when considering the variance explained and number of items by factor (Factor 1 $r = -.239$, $n = 26$; Factor 2 $r = -.171$, $n = 23$; Factor 3 $r = -.073$, $n = 14$; Factor 4 $r = -.164$, $n = 11$). These correlations are small but consistently negative and may speak to the retention of items that are not sufficiently strong in the factor solution (i.e., items that add more unexplained than explained variance to the model). Given that authors likely often reported the factor level variance explained prior to rotation, the correlations for the first factors may even be somewhat attenuated.

These issues are worthy of further empirical investigation. If analysts are not extracting the correct number of factors, subsequent results can be adversely affected. If analysts' instruments are not yielding scores with factorial "simple structure" (Thurstone, 1935) or include items that add more unexplained than explained variance, the construct validity of scores may be questionable. Alternatively, perhaps a more realistic expectation of the amount of variance that should be explained in a factor analysis is warranted.

Table 2 presents frequency counts and percentages of articles reporting various elements of factor analytic information. Table 2 presents both overall frequencies as well as those for each journal examined. For the sake of brevity, only the overall results will be summarized here. However, it should be noted that in general, the outcomes for the individual journals were similar to the overall results, with the marked exception of article type. Article type varied considerably because of the different objectives, both substantive and measurement, of the journals examined.

Careful examination of Table 2 highlights many of the typical decisions made by researchers when conducting EFAs. Careful review also reveals some egregious errors concerning appropriate reporting practice. For example, the majority (65.0%) of authors failed to note what matrix of association they analyzed. Authors also failed to indicate their factor extraction method on eight occasions (13.3%). Among those reporting the extraction method used, most (56.7%) used PCA, which is the default option in most statistical packages.

Regarding strategies used to determine the number of factors to retain, the EV > 1 rule was most common (56.7%). Interestingly, the EV > 1 rule is also the default in most statistical packages, and its frequency of use mirrors that for PCA. The scree test was frequently used (35.0%) as well. Largely, the other rules were ignored (or at least not reported, and so assumed ignored) by authors. For 10 uses of EFA, the authors

(text continues on page 407)

Table 2
Frequencies and Percentage of Articles Reporting EFA Information

	EPM (n = 15)		*JEP* (n = 15)		*PA* (n = 15)		*P&ID* (n = 15)		Overall (N = 60)	
Variable	n	%	n	%	n	%	n	%	n	%
Article type										
Measurement	15	100.0	—	—	14	93.3	7	46.7	36	60.0
Substantive	—	—	15	100.0	1	6.7	8	53.3	24	40.0
Level										
First order	15	100.0	14	93.3	15	100.0	14	93.3	58	96.7
Higher order	—	—	1	6.7	—	—	1	6.7	2	3.3
Matrix analyzed										
Correlation	5	33.3	6	40.0	3	20.0	5	33.3	19	31.7
Variance/covariance	—	—	—	—	—	—	2	13.3	2	3.3
Not reported	10	66.7	9	60.0	12	80.0	8	53.3	39	65.0
Extraction method										
Principal components	11	73.3	6	40.0	10	66.7	7	46.7	34	56.7
Principal axis	1	6.7	5	33.3	5	33.3	2	13.3	13	21.7
Other	—	—	1	6.7	—	—	4	26.7	5	8.3
Not reported	3	20.0	3	20.0	—	—	2	13.3	8	13.3
Strategies used for factor retention										
EV > 1	9	60.0	9	60.0	7	46.7	9	60.0	34	56.7
Scree plot	6	40.0	1	6.7	9	60.0	5	33.3	21	35.0
Minimum average partial	—	—	—	—	—	—	—	—	0	0.0
Parallel analysis	1	6.7	1	6.7	1	6.7	1	6.7	4	6.7
Bartlett's chi-square	—	—	—	—	—	—	—	—	0	0.0
No. set a priori	3	20.0	2	13.3	2	13.3	3	20.0	10	16.7
Other	—	—	1	6.7	2	13.3	4	26.7	7	11.7
General rotation strategy										
Orthogonal	6	40.0	9	60.0	7	46.7	11	73.3	33	55.0
Oblique	7	46.7	5	33.3	8	53.3	3	20.0	23	38.3
No rotation used	—	—	1	6.7	—	—	—	—	1	1.7

Not reported	2	13.3	—	—	—	—	1	6.7	3	5.0
Rotation justification presented										
Yes	6	40.0	4	26.7	8	53.3	5	33.3	23	38.3
No	9	60.0	10[a]	66.7	7	46.7	10	66.7	37	61.7
Rotation type										
Varimax	6	40.0	8	53.3	6	40.0	11	73.3	31	51.7
Oblimin	3	20.0	2	13.3	5	33.3	3	20.0	13	21.7
Delta value given										
Yes (delta = 0)	1	33.3	—	—	2	40.0	—	—	3	23.1
No	2	66.7	2	100.0	3	60.0	3	100.0	10	76.9
Promax	3	20.0	1	6.7	—	—	—	—	4	6.7
Pivot given										
Yes	—	—	—	—	—	—	—	—	0	0.0
No	3	100.0	1	100.0	—	—	—	—	4	100.0
Procrustes	—	—	—	—	1	6.7	—	—	1	1.7
Other	1	6.7	2	13.3	—	—	—	—	3	5.0
Not reported	2	13.3	1[a]	6.7	3	20.0	1	6.7	8	13.4
If oblique, coefficients reported (n = 23 total)										
Factor pattern only	3	42.9	1	20.0	4	50.0	3	100.0	11	47.8
Factor structure only	1	14.3	1	20.0	2	25.0	—	—	4	17.4
Both	1	14.3	—	—	—	—	—	—	1	4.3
Cannot tell	1	14.3	1	20.0	—	—	—	—	2	8.7
None reported	1	14.3	2	40.0	2	25.0	—	—	5	21.7
Reported communalities										
Yes	2	13.3	2	13.3	2	13.3	4	26.7	10	16.7
No	13	86.7	13	86.7	13	86.7	11	73.4	50	83.4
Reported variance explained/factor										
Yes	5	33.3	3	20.0	9	60.0	11	73.3	22	36.7
No	10	66.7	12	80.0	6	40.0	4	26.7	38	63.3
Named factors other than variable name										
Yes	10	66.7	9	60.0	15	100.0	11	73.3	45	75.0

(continued)

Table 2 (continued)

	EPM (n = 15)		*JEP* (n = 15)		*PA* (n = 15)		*P&ID* (n = 15)		Overall (N = 60)	
Variable	n	%	n	%	n	%	n	%	n	%
No	5	33.3	6	40.0	—	—	4	26.7	15	25.0
Reported EV for factors retained										
Yes	6	40.0	6	40.0	7	46.7	10	66.7	29	48.3
No	9	60.0	9	60.0	8	53.3	5	33.3	31	51.7
Reported EV of at least one factor not retained										
Yes	—	—	—	—	2	13.3	1	6.7	3	5.0
No	15	100.0	15	100.0	13	86.7	14	93.3	57	95.0
Initial EV interpreted as applying postrotation										
Yes	—	—	—	—	3	20.0	7	46.7	10	16.7
No	6	40.0	6	40.0	3	20.0	3	20.0	19	31.7
No reference	9	60.0	9	60.0	9	60.0	5	33.3	31	51.7
CFA warranted										
Yes, not new measure	7	46.7	3	20.0	4	26.7	6	40.0	20	33.3
No, new measure	5	33.3	11	78.3	8	53.3	9	60.0	33	55.0
Yes, but both EFA and CFA done	3	20.0	1	6.7	3	20.0	—	—	7	11.7
If CFA warranted, reasons given for not using CFA (n = 20 total)										
Sample size	—	—	—	—	—	—	—	—	0	0.0
No strong theory	—	—	—	—	—	—	—	—	0	0.0
Other	—	—	—	—	—	—	—	—	0	0.0
Not addressed	7	100.0	3	100.0	4	100.0	6	100.0	20	100.0

Note: *EPM* = *Educational and Psychological Measurement*; *JEP* = *Journal of Educational Psychology*; *PA* = *Psychological Assessment*; *P&ID* = *Personality and Individual Differences*; EV = eigenvalue; CFA = confirmatory factor analysis; EFA = exploratory factor analysis.
a. The n does not sum to 15 because only one factor was extracted in one case. No rotation was necessary or possible.

noted that they set the number of factors to extract based on a priori theory. Daniel (1989) and Kieffer (1999) suggested that this approach is generally less than optimal, and a CFA may be more appropriate in these circumstances.

It is important to note, however, that an EFA may indeed be warranted during instrument development, even when theoretical expectations are present regarding the number of factors. For example, theory often drives item development, and these items are often subsequently assessed with EFA to help refine the assessment. Nevertheless, such preliminary hypotheses about the number of factors should be considered in conjunction with other factor retention rules at the exploratory stage, given that a priori expectations may in fact be incorrect.

Although Zwick and Velicer (1986) demonstrated that minimum average partial and parallel analysis were among the most accurate methods for determining the number of factors, most authors failed to use either of these methods. Minimum average partial was never used, and parallel analysis (Turner, 1998) was used on four occasions (6.7%). Given the problems with the EV > 1 rule, and less so with the scree test, these findings are troublesome and call into question whether the authors extracted the correct number of factors from their matrices. It is in cases such as these that independent evaluation of results is critical; however, few articles reported enough information to allow for such an investigation.

Furthermore, despite Thompson and Daniel's (1996) recommendation that "the simultaneous use of multiple decision rules is appropriate and often desirable" (p. 200), most authors in the current study ($n = 33$, 55.0%) only used (or at least only reported using) one rule. Of these, 18 invoked the EV > 1 rule, 6 used the scree test, 2 used parallel analysis, and 7 made decisions based on a priori theory. Two decision rules were explicitly considered in 11 EFAs, and three rules were used in 7 EFAs. No authors reported using more than three rules. Unfortunately, authors of 9 EFAs (15%) failed to give any indication of how they determined the number of factors to extract.

Regarding factor rotation, orthogonal rotations were most common (55.0%), although rotation strategy was not explicitly justified in 61.7% of the EFAs. Varimax was the most commonly used specific method (51.7%). The most common oblique strategy was oblimin (21.7%), although the exact delta value used was not given in almost all of these cases ($n = 10$). Gorsuch (1983) discussed the potential differences from using varied delta values. When delta was reported ($n = 3$), it was always 0, the default in most statistical packages.

Thirteen percent of cases did not report their specific rotation strategy, and 3 failed to even indicate whether the rotation was orthogonal or oblique. Again, this lack of information severely limits external evaluation of others' work. The reader is left to accept the authors' findings on faith, a noble virtue in some contexts but not in EFA.

Furthermore, when oblique solutions were used ($n = 23$), only 5 EFAs included either the factor structure matrix ($n = 4$) or both the factor pattern matrix and the structure matrix ($n = 1$). The rest erroneously reported only the factor pattern matrix ($n = 11$, which is insufficient for the reasons noted previously), did not report any of the matrices ($n = 5$), or presented the matrix ambiguously so as to prevent the reader from knowing what matrix was given ($n = 2$). It can only be assumed that authors used the

Table 3
Variance Explained and Number of Items for Reported Factors

	% Variance Explained			No. of Items				
Factor	*n*	*M*	*SD*	*n*	*M*	*SD*	Minimum	Maximum
Reported before rotation								
1	17	31.18	11.06	24	7.54	3.41	3	17
2	16	10.60	5.86	23	5.09	2.70	2	12
3	9	7.52	3.54	15	4.73	2.34	2	9
4	7	7.42	2.01	12	5.67	3.26	2	13
5	1	5.20	—	4	7.00	1.83	5	9
6				2	10.50	0.71	10	11
7				1	7.00	—	7	7
Reported after rotation								
1	4	29.75	30.56	6	7.67	3.67	4	14
2	4	11.45	1.66	6	6.67	3.78	1	10
3	3	10.00	1.05	5	5.60	2.88	2	8
4	3	8.00	0.10	4	5.00	2.94	1	8
5	1	10.00	—	2	6.50	0.71	6	7

matrices reported to make substantive interpretations of the factor structure. When only the factor pattern matrix is consulted in oblique solutions, incorrect interpretations are very possible (Gorsuch, 1983; Kieffer, 1999; Thompson & Daniel, 1996). Structure coefficients are also almost always necessary for interpretation when factors are correlated.

Additional errors of omission included failure to report communality coefficients (83.4%), variance explained for each factor (63.3%), and EVs for each factor prior to rotation (51.7%). We would also suggest that external evaluation would be facilitated by reporting the EV for at least the first factor not retained. This EV would be particularly relevant when only the EV > 1 rule is used for extraction. Only 5.0% of the EFAs reported this information.

For interpretation clarity, Thompson and Daniel (1996) suggested that "factors should be given names that do not invoke the labels of observed variables because the latent constructs are not observed variables themselves" (p. 202). Seventy-five percent of the EFAs met this expectation. Another aspect of factor interpretation involves how many and how strongly observed variables weight on a given factor. At least two variables are necessary to define a factor—otherwise the factor would be little more than the observed variable itself. Although multiple items were used to define factors in most all cases, 6 of the EFAs involved factors that were defined by only one variable, which seems to contradict the basic idea of a factor as a latent construct.

Table 3 presents descriptive information regarding the variance explained by extracted factors and the number of salient items per factor. The data are presented based on whether the authors reported the variance explained before or (appropriately) after rotation. The average number of salient items for a given factor was approximately six.

Finally, we also examined whether CFA was warranted as a potentially more appropriate analysis when the authors held a priori expectations concerning the factor structure. In general, we considered a priori theory tenable when the instrument was not new and when the authors had knowledge of the factor structure of scores from a previous administration of the instrument. In these cases, it can be argued that CFA may be a preferred method given its ability to falsify theoretical expectations, whereas EFA may be more appropriate during instrument development. It is recognized that EFA can be used in confirmatory ways, but CFA represents a more direct approach. Furthermore, as noted above, EFA can indeed be useful even when a priori theory is present. The utility of EFA or CFA depends in large part to the strength of the prior theory, which of course is more of a continuum than an absolute. Accordingly, the current definition for CFA use seemed plausible for the purposes of this review.

In our sample, CFA was warranted but not used in one third of the cases. Some authors (11.7%) conducted an EFA when they had theoretical expectations but then appropriately followed up the EFA with a CFA on an independent sample. Although CFA may not be tenable in some instances (e.g., small sample size), this finding reflects a tendency to underuse CFA.

In CFA, multiple models can be pitted against each other in an attempt to falsify the theoretical constructs that are tested. This falsification potential is fundamental to construct validity and theory development. As noted by Thompson and Daniel (1996),

> CFA can readily be used to test rival models and to quantify the fit of each rival model. Testing rival models is usually essential because multiple models may fit the same data. Of course, finding that a single model fits data well, whereas other plausible models do not, does not "prove" the model, since untested models may fit even better. However, testing multiple plausible models does yield stronger evidence regarding validity. (p. 204)

Recommendations for Practice

Based on the results obtained in the present study and the pleas of numerous researchers (cf. Comrey, 1978; Fabrigar et al., 1999; Gorsuch, 1983; Henson et al., 2001; Kline, 1994; Russell, 2002; Thompson & Daniel, 1996; Tinsley & Tinsley, 1987; Weiss, 1971), we suggest the following recommendations for practice when conducting and reporting an EFA. In general, sufficient information should be presented to allow external evaluation of the analysis, and all analytic decisions should be explicitly noted. Unfortunately, these expectations were often unmet in the articles examined here.

1. When prior theory exists regarding the structure of the data, CFA should be considered as an alternative to EFA.
2. Always report which matrix of association was analyzed and the method of factor extraction used. Furthermore, the actual matrix of association should be reported (or made available on request) to allow others to replicate the analysis.

3. Use and report multiple criteria when determining the number of factors to retain. Avoid overdependence on the EV > 1 rule. Parallel analysis and minimum average partial are grossly underused in published research and should be employed with greater frequency, given their utility (cf. Zwick & Velicer, 1986). O'Connor (2000) and Thompson and Daniel (1996) provided programs to conduct parallel analysis for interested readers. We also suggest that authors report the EV for the first factor not retained.
4. Explicitly indicate which specific rotation strategy was used (e.g., varimax, promax). Furthermore, explicitly justify why an orthogonal or an oblique solution was selected. In general, as is the case throughout the general linear model, because oblique rotation requires the estimation of more parameters, an oblique structure will usually fit sample data better than will an orthogonal rotation. However, as Hetzel (1996) noted,

    > some researchers have argued that, all things being equal, orthogonal solutions are desirable. Since the factor pattern and the factor structure matrices are identical, and the factor correlation matrix is an identity matrix, fewer parameter matrices are estimated. In theory, the resulting parsimony should lead to more replicable results. (p. 194)

    In either case, the decision should be clearly justified and not made by default. Authors should first obliquely rotate results, examine the factor intercorrelations (and report them), and then make judgments about whether an orthogonal or oblique rotation is most warranted.
5. Always report the full factor pattern/structure matrix. All factored items should be included. This information is needed to allow (a) external evaluation of analytic decisions, (b) others to rotate reported results to alternative rotation criteria, and (c) meta-analytic investigation of factor structure invariance across studies. When oblique solutions are used, always report and interpret both the factor pattern and factor structure matrices.

    Table 4 illustrates a recommended reporting method when presenting orthogonal factor pattern/structure coefficients. Although this table does not list such information as the justification for the rotation strategy, it does present all pertinent information concerning results from the EFA. Although this table alone would not allow an external researcher to reproduce a presented study, it should help readers understand the general design of the study and relevant outcome information concerning the results.
6. Always report communalities, the total variance explained by the factors, initial EVs, and the variance explained by each factor after rotation or final traces (i.e., the transformed EV variance-accounted-for statistic after rotation).
7. Never name a factor with the label used for an observed variable. Such practice is potentially confusing and does not honor the fact that the factor is a latent, unobserved variable. In addition, do not define a factor with only one item. Sufficient component saturation is needed to warrant factor interpretation and to assume some level of replicability.

Conclusion

Appropriate use of EFA necessitates thoughtful researcher judgment concerning a number of analytic decisions. The present article reviewed some of the fundamental decisions necessary to conduct an EFA and examined reporting practice in published research. Largely, the information presented in published applications of EFA results tended to be insufficient to allow external verification of the EFA results and re-

Table 4
Heuristic Factor Pattern/Structure Matrix Rotated to the Varimax Criterion

Variable	Factor 1: Mechanical	Factor 2: Spatial	Factor 3: Verbal	h^2
X1	*.685*	.133	.168	.515
X2	.005	.070	*.832*	.697
X3	*.625*	.280	.032	.470
X4	.101	*.688*	–.110	.496
X5	.035	.003	*.850*	.724
X6	*.489*	.358	.252	.431
X7	*.822*	.085	.008	.683
X8	.006	–.002	*.780*	.608
X9	.285	*.589*	.056	.431
X10	.100	*.785*	.021	.627
Trace	1.841	1.673	2.132	5.646
% of variance	18.4	16.7	21.3	56.4

Note: Coefficients greater than |.40| are italicized and retained for that factor. Percentage variance is postrotation; because here there were 10 measured variables, percentage of variance is trace divided by 10 times 100 (or trace times 10). The eigenvalue of the fourth, unretained factor was 1.102. h^2 = communality coefficient.

searcher decisions. In addition, the results suggest that researchers often simply use the default options in common statistical packages, which may lead to errant results. For example, when determining the number of factors to retain, the default option (usually EV > 1) is among the weakest methods available.

Multiple deficits in reported information were noted. Several recommendations for improved EFA reporting practice were also given, including the overall recommendation of providing sufficient information to allow external verification of one's EFA results and decisions. Historically, factor analytic techniques have been very useful in theory development and assessing construct validity (cf. Nunnally, 1978, p. 112). The future value of EFA would be enhanced by (a) careful consideration of the choices made in the analysis and (b) reporting more complete information in published research.

References

*References marked with an asterisk indicate studies included in the current review.

*Arrindell, W. A., Heesink, J., & Feij, J. A. (1999). The Satisfaction With Life Scale (SWLS): Appraisal with 1700 healthy young adults in the Netherlands. *Personality and Individual Differences*, *26*, 815-826.

Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of a structural relations model. *Multivariate Behavioral Research*, *16*, 437-454.

*Ball, S. A., Tennen, H., & Kranzler, H. R. (1999). Factor replicability and validity of the Temperament and Character Inventory in substance-dependent patients. *Psychological Assessment*, *11*, 514-524.

Bandalos, B. (1996). Confirmatory factor analysis. In J. Stevens (Ed.), *Applied multivariate statistics for the social sciences* (3rd ed., pp. 389-420). Mahwah, NJ: Lawrence Erlbaum.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, *3*, 77-85.
Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology*, *4*, 1-2.
*Bowen, C. W. (1999). Development and score validation of a Chemistry Laboratory Anxiety Instrument (CLAI) for college chemistry students. *Educational and Psychological Measurement*, *59*, 171-185.
*Brady, K. L., & Eisler, R. M. (1999). Sex and gender in the college classroom: A quantitative analysis of faculty-student interactions and perceptions. *Journal of Educational Psychology*, *91*, 127-145.
*Budaev, S. V. (1999). Sex differences in the Big Five personality factors: Testing an evolutionary hypothesis. *Personality and Individual Differences*, *26*, 801-813.
*Butler, R. (1998). Determinants of help seeking: Relations between perceived reasons for classroom help-avoidance and help-seeking behaviors in an experimental context. *Journal of Educational Psychology*, *90*, 630-643.
Campbell, T. (1996). Investigating structures underlying relationships when variables are not the focus: Q-technique and other techniques. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 207-218). Greenwich, CT: JAI.
*Carter, M. M., Miller, O., Jr., Sbrocco, T., Suchday, S., & Lewis, E. L. (1999). Factor structure of the Anxiety Sensitivity Index among African American college students. *Psychological Assessment*, *11*, 525-533.
*Cascardi, M., Avery-Leaf, S., O'Leary, K. D., & Smith Slep, A. M. (1999). Factor structure and convergent validity of the Conflict Tactics Scale in high school students. *Psychological Assessment*, *11*, 546-555.
Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245-276.
Cattell, R. B., & Jaspers, J. (1967). A general plasmode for factor analytic exercises and research. *Multivariate Behavioral Research Monographs*, *3*, 1-212.
*Clark, R. (1999). The Parent-Child Early Relational Assessment: A factorial validity study. *Educational and Psychological Measurement*, *59*, 821-846.
Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin*, *70*, 426-443.
Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, *46*, 648-659.
Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
*Coster, W. J., Mancini, M. C., & Ludlow, L. H. (1999). Factor structure of the School Function Assessment. *Educational and Psychological Measurement*, *59*, 665-677.
Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement*, *61*, 229-248.
Cronkhite, G., & Liska, J. R. (1980). The judgment of communicant acceptability. In M. R. Roloff & G. R. Miller (Eds.), *Persuasion: New directions in theory and research* (pp. 101-139). Beverly Hills, CA: Sage.
*Dag, I. (1999). The relationships among paranormal beliefs, locus of control and psychopathology in a Turkish college sample. *Personality and Individual Differences*, *26*, 723-737.
Daniel, L. G. (1989, November). *Comparisons of exploratory and confirmatory factor analysis*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED314447)
*Davis, C., Tang, C. S., Chan, S., & Noel, B. (1999). The development and validation of the International AIDS Questionnaire–Chinese version (IAQ-C). *Educational and Psychological Measurement*, *59*, 481-491.
*Endler, N. S., Parker, J. D. A., & Summerfeldt, L. J. (1999). Coping with health problems: Developing a reliable and valid multidimensional measure. *Psychological Assessment*, *10*, 195-205.
Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272-299.
Fan, X. (1996). Canonical correlation analysis as a general analytic model. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 71-94). Greenwich, CT: JAI.

Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common? *Structural Equation Modeling*, *4*, 65-79.
*Feltz, D. L., Chase, M. A., Moritz, S. E., & Sullivan, P. J. (1999). A conceptual model of coaching efficacy: Preliminary investigation and instrument development. *Journal of Educational Psychology*, *91*, 765-776.
*Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. M. (1999). The Posttraumatic Cognitions Inventory (PTCI): Development and validation. *Psychological Assessment*, *11*, 303-314.
*Forsterlee, R., & Ho, R. (1999). An examination of the short form of the Need for Cognition Scale applied in an Australian sample. *Educational and Psychological Measurement*, *59*, 471-480.
*Friedman, I. A. (1999). Teacher-perceived work autonomy: The concept and its measurement. *Educational and Psychological Measurement*, *59*, 58-76.
*Furnham, A., Fong, G., & Martin, N. (1999). Sex and cross-cultural differences in the estimated multifaceted intelligence quotient score for self, parents and siblings. *Personality and Individual Differences*, *26*, 1025-1034.
*Goldberg, L. R. (1999). The Curious Experiences Survey, a revised version of the Dissociative Experiences Scale: Factor structure, reliability, and relations to demographic and personality variables. *Psychological Assessment*, *11*, 134-145.
Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
*Hagemann, D., Naumann, E., Maier, S., Becker, G., Alexander, L., & Dieter, B. (1999). The assessment of affective reactivity using films: Validity, reliability and sex differences. *Personality and Individual Differences*, *26*, 627-639.
Hendrick, C., & Hendrick, S. (1986). A theory and method of love. *Journal of Personality and Social Psychology*, *50*, 392-402.
Henson, R. K. (2000). Demystifying parametric analyses: Illustrating canonical correlation as the multivariate general linear model. *Multiple Linear Regression Viewpoints*, *26*(1), 11-19.
Henson, R. K. (2002, April). *The logic and interpretation of structure coefficients in multivariate general linear model analyses*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED467381)
Henson, R. K., Capraro, R. M., & Capraro, M. M. (2004). Reporting practices and use of exploratory factor analyses in educational research journals: Errors and explanation. *Research in the Schools, 11*(2), 61-72.
Henson, R. K., & Smith, A. D. (2000). State of the art in statistical significance and effect size reporting: A review of the APA Task Force Report and current trends. *Journal of Research and Development in Education*, *33*(4), 284-294.
*Hess, C. W., & Wagner, B. T. (1999). Factor structure of the Rahim Leader Power Inventory (RLPI) with clinical female student supervisees. *Educational and Psychological Measurement*, *59*, 1004-1015.
Hetzel, R. D. (1996). A primer on factor analysis with comments on patterns of practice and reporting. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 175-206). Greenwich, CT: JAI.
Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179-185.
Humphreys, L. G. (1964). Number of cases and number of factors: An example where N is very large. *Educational and Psychological Measurement*, *24*, 457.
*Johnson, W. L. (1999). A primary- and second-order analysis of the Organizational Identification Questionnaire. *Educational and Psychological Measurement*, *59*, 159-170.
*Johnson, W. L., Johnson, A. M., Kranch, D. A., & Zimmerman, K. J. (1999). The development of a university version of the Charles F. Kettering Climate Scale. *Educational and Psychological Measurement*, *59*, 336-350.
Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183-202.
Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141-151.
*Kardash, C. M, & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology*, *88*, 260-271.

Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach*. New York: Holt, Rinehart & Winston.

Kieffer, K. M. (1999). An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. *Research in the Schools*, *6*, 75-92.

*Klecker, B. M., & Loadman, W. E. (1998). Another look at the dimensionality of the School Participant Empowerment Scale. *Educational and Psychological Measurement*, *58*, 944-954.

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. *Psychological Bulletin*, *85*, 410-416.

*Lajunen, T., & Hanna, R. S. (1999). Is the EPQ Lie Scale bidimensional? Validation study of the structure of the EPQ Lie Scale among Finnish and Turkish university students. *Personality and Individual Differences*, *26*, 657-664.

*Laurent, J., Catanzaro, S. J., Joiner, T. E., Jr., Rudolph, K. D., Potter, K. I., Lambert, S., et al. (1999). A measure of positive and negative affect for children: Scale development and preliminary validation. *Psychological Assessment*, *11*, 326-338.

Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. *Psychometrika*, *33*, 37-71.

*Loo, R., & Thorpe, K. (1999). A psychometric investigation of scores on the Watson-Glaser critical thinking appraisal new form S. *Educational and Psychological Measurement*, *59*, 995-1003.

*Lovejoy, M. C., Weis, R., O'Hare, E., & Rubin, E. (1999). Development and initial validation of the Parent Behavior Inventory. *Psychological Assessment*, *11*, 534-545.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84-99.

Mote, T. A. (1970). An artifact of the rotation of too few factors: Study orientation vs. trait anxiety. *Revista Interamericana de Psicologia*, *37*, 61-91.

Mulaik, S. A. (1987). A brief history of the philosophical foundations of exploratory factor analysis. *Multivariate Behavioral Research*, *22*, 267-305.

*Newstead, S. E., Franklyn-Stokes, A., & Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology*, *88*, 229-241.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396-402.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, *6*, 559-572.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.

*Ponterotto, J. G., Baluch, S., Grieg, T., & Rivera, L. (1998). Development and initial score validation of the Teacher Multicultural Attitude Survey. *Educational and Psychological Measurement*, *58*, 1002-1016.

*Riordan, C. M., & Weatherly, E. W. (1999). Defining and measuring employees' identification with their work groups. *Educational and Psychological Measurement*, *59*, 310-324.

*Rolfhus, E. L., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: Knowledge, intelligence, and related traits. *Journal of Educational Psychology*, *91*, 511-526.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, *28*, 1626-1646.

*Sadoski, M., Kealy, W. A., Goetz, E. T., & Paivio, A. (1997). Concreteness and imagery effects in the written composition of definitions. *Journal of Educational Psychology*, *89*, 518-526.

*Schraw, G., & Nietfeld, J. (1998). A further test of the general monitoring skill hypothesis. *Journal of Educational Psychology*, *90*, 236-248.

*Scott, V. B., Jr., & McIntosh, W. D. (1999). The development of a trait measure of ruminative thought. *Personality and Individual Differences*, *26*, 1045-1056.

*Shafer, A. B. (1999). Factor analyses of Big Five markers with the Comrey Personality Scales and the Howarth Personality Tests. *Personality and Individual Differences*, *26*, 857-872.

*Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology*, *89*, 71-81.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148-154.

*Sparks, R. L., Ganschow, L., & Patton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology*, *87*, 638-655.

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

*Stern, P. C., Dietz, T., & Guagnano, G. A. (1998). A brief inventory of values. *Educational and Psychological Measurement*, *58*, 984-1001.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

*Stipek, D., & Gralinski, J. H. (1996). Children's beliefs about intelligence and school performance. *Journal of Educational Psychology*, *88*, 397-407.

*Swanson, H. L., & Alexander, J. E. (1997). Cognitive processes as predictors of word recognition and reading comprehension in learning-disabled and skilled readers: Revisiting the specificity hypothesis. *Journal of Educational Psychology*, *89*, 128-158.

*Swanson, H. L., Mink, J., & Bocian, K. M. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: More alike than different? *Journal of Educational Psychology*, *91*, 321-333.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins.

Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation*. Beverly Hills, CA: Sage.

Thompson, B. (1988). Program FACSTRAP: A program that computes bootstrap estimates of factor structure. *Educational and Psychological Measurement*, *48*, 681-686.

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development*, *24*, 80-95.

Thompson, B. (1992). A partial test distribution for cosines among factors across samples. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 2, pp. 81-97). Greenwich, CT: JAI.

Thompson, B. (1999). If statistical significance tests are broken/misused, what practices should supplement or replace them? *Theory and Psychology*, *9*, 165-181.

Thompson, B. (2000a). Canonical correlation analysis. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 285-316). Washington, DC: American Psychological Association.

Thompson, B. (2000b). Q-technique factor analysis: One variation on the two-mode factor analysis of variables. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 207-226). Washington, DC: American Psychological Association.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis*. Washington, DC: American Psychological Association.

Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. *Educational and Psychological Measurement*, *45*, 203-209.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, *56*, 197-208.

*Thorkildsen, T. A., & Nicholls, J. G. (1998). Fifth graders' achievement orientations and beliefs: Individual classroom differences. *Journal of Educational Psychology*, *90*, 179-201.

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

Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*, 414-424.

*Tropp, L. R., Erkut, S., Coll, C. G., Alarcón, O., & Vázquez García, H. A. (1999). Psychological acculturation: Development of a new measure for Puerto Ricans on the U.S. mainland. *Educational and Psychological Measurement*, *59*, 351-367.

Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement*, *58*, 541-568.

*Van Gerwen, L. J., Spinhoven, P., Dyck, R. V., & Diekstra, R. F. W. (1999). Construction and psychometric characteristics of two self-report questionnaires for the assessment of fear of flying. *Psychological Assessment*, *11*, 146-158.

*Van Lankveld, J. J. D. M., & Ter Kuile, M. M. (1999). The Golombok Rust Inventory of Sexual Satisfaction (GRISS): Predictive validity and construct validity in a Dutch population. *Personality and Individual Differences*, *26*, 1005-1023.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321-327.

Weiss, D. J. (1971). Further considerations in applications of factor analysis. *Journal of Counseling Psychology*, *18*, 85-92.

*Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, *89*, 420-432.

*Yang, J., McCrae, R. R., Costa, P. T., Jr., Dai, X., Yao, S., Cai, T., et al. (1999). Cross-cultural personality assessment in psychiatric populations: The NEO-PI-R in the People's Republic of China. *Psychological Assessment*, *11*, 359-368.

Yeomans, K. A., & Golder, P. A. (1982). The Guttman-Kaiser criterion as a predictor of the number of common factors. *Statistician*, *31*, 221-229.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, *17*, 253-269.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432-442.