International Journal of Human-Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: http://www.tandfonline.com/loi/hihc20

A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve?

Matt C. Howard

Intl. Journal of Human–Computer Interaction, 32: 51–62, 2016

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447318.2015.1087664

A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve?

Matt C. Howard
Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania

Authors within the fields of cyberpsychology and human−computer interaction have demonstrated a particular interest in measurement and scale creation, and exploratory factor analysis (EFA) is an extremely important statistical method for these areas of research. Unfortunately, EFA requires several statistical and methodological decisions to which the best choices are often unclear. The current article reviews five primary decisions and provides direct suggestions for best practices. These decisions are (a) the data inspection techniques, (b) the factor analytic method, (c) the factor retention method, (d) the factor rotation method, and (e) the factor loading cutoff. Then the article reviews authors' choices for these five EFA decisions in every relevant article within seven cyberpsychology and/or human–computer interaction journals. The results demonstrate that authors do not employ the recommended best practices for most decisions. Particularly, most authors do not inspect their data for violations of assumptions, apply inappropriate factor analytic methods, utilize outdated factor retention methods, and omit the justification for their factor rotation methods. Further, many authors omit altogether their EFA decisions. To rectify these concerns, the current article provides a step-by-step guide and checklist that authors can reference to ensure the use of recommended best practices. Together, the current article identifies concerns with current research and provides direct solutions to these concerns.

Ensuring proper scale creation and evaluation methods is particularly important for the study of cyberpsychology and human–computer interactions for several reasons. First, these research areas are often interested in the relationship between technology and human cognitions, affect, and personality. These human dynamics are conceptualized as unobservable latent constructs that are identified and measured through scales (Hinkin, 1995, 1998; Tabachnick & Fidell, 2007). Second, these fields are relatively new. The psychometric properties and validity of scales measuring many constructs of interest have yet to be properly analyzed, but authors require sound measures to draw accurate research inferences. For these reasons and others, scholars within these fields have demonstrated a great interest in proper scale creation and evaluation methods (Hamari & Koivisto, 2014; Howard & Jayne, 2015; Lam & Li, 2013).

Although the scale development and evaluation process consists of several steps, among the most important is identifying a theoretically and psychometrically sound factor structure (Comrey & Lee, 1992; Henson & Roberts, 2006). Providing relevant theory differs for each scale, but identifying an initial factor structure through exploratory factor analysis (EFA) is often mandatory (Costello & Osborne, 2005; Fabrigar, Wegener, MacCallum, & Strahan, 1999). Performing an EFA requires several statistical and methodological decisions. Unfortunately, the correct choices for these decisions are often unclear, causing scholars to obtain inaccurate EFA results (Gorsuch, 1997; Henson & Roberts, 2006). Due to the importance of the scale creation process to cyberpsychology and human–computer interaction scholarship, the current article has several objectives related to identifying proper EFA methods.

First, a review of proper EFA methods is provided, and each required decision when performing an EFA is clearly stated. Second, the current article catalogues the statistical and methodological decisions of all authors applying EFA within seven journals, resulting in clear inferences about the current practices of scale creation and evaluation within cyberpsychology and human–computer interaction research. Third, direct suggestions are given for future researchers to overcome the common concerns discovered through reviewing prior EFA decisions. Fourth, implications for future research and practice are noted. Together, the current article may greatly improve the future of scale development and evaluation through drawing attention to common difficulties when performing EFA.

Further, it should be noted that the current article is not meant to be an overview of cutting-edge EFA practices, such as Bayesian factor analysis and exploratory structural equation modeling. Many of these methods can be performed only with a knowledge of specialized software, and they are currently outside the scope of many authors who apply EFA in cyberpsychology and human–computer interaction studies.

Address correspondence to Matt C. Howard, Department of Psychology, 142 Moore Building, The Pennsylvania State University, University Park, PA 16802, USA. E-mail: mch248@psu.edu

Instead, the current article focuses on methods that can be easily applied by anyone with a general knowledge of SPSS or SAS, which is assumed to be most researchers. With these aspects noted, an overview of best-practices and researchers' EFA decisions is presented.

1. EFA DECISIONS

When performing an EFA, several statistical and methodological decisions are required. These are (a) the data inspection techniques, (b) the factor analytic method, (c) the factor retention method, (d) the factor rotation method, and (e) the factor loading cutoff. Each of these should be planned before performing the EFA, as the decisions should be based upon prior theory and methodological logic. If an author makes these decisions during the scale creation or evaluation process, such as through testing multiple methods, they may inadvertently increase their familywise error rate and risk committing a Type I error (Benjamini & Hochberg, 1995). More important, they also risk their decisions being chosen by the preferred results rather than the *correct* results, which is a clear violation of scientific integrity.

Further, it should be noted that the current review is not meant to be entirely comprehensive. Each EFA decision itself could constitute a book. Instead, the current review is meant to be accessible to most researchers and provide general EFA advice. Scholars should stay alert toward additional methods that may be incorporated within the EFA process. Psychometricians consistently discover methodological and statistical improvements, and best practices are constantly changing. With this in mind, each decision is reviewed next.

1.1. Data Inspection Techniques

With any analysis, authors should check their data for violations of statistical assumptions, and EFA is no different. The two most popular data inspection techniques for EFA are Bartlett's test of sphericity (Bartlett, 1950; Dziuban & Shirkey, 1974) and the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy (Dziuban & Shirkey, 1974; Kaiser, 1970). Both of these methods test whether sufficiently large relationships exist within the data set of interest to perform EFA.

Bartlett's test of sphericity checks whether the observed correlation matrix is an identity matrix, which holds the property of having all off-diagonal values of zero (Tobias & Carlson, 1969). Given that factor analysis explains the relationships of variables, a complete lack of relationships within a data set (i.e., an identity matrix) prevents EFA from being performed. If Bartlett's test of sphericity is significant, the results indicate that the data are not an identity matrix and appropriate for EFA. Although this test is successful in checking for violations of EFA assumptions, authors have noted that virtually all data sets are significantly different from an identity matrix, and the Bartlett's test of sphericity is rarely nonsignificant (Dziuban & Harris, 1973; Dziuban & Shirkey, 1974). Nevertheless, the test may detect problematic data sets, and it should be performed prior to EFA.

Alternatively, the KMO Measure of Sampling Adequacy is an indicator of common variance within a data set, which indicates that latent factors may be present and EFA may be performed (Dziuban & Shirkey, 1974; Kaiser, 1970). In general, the results of Bartlett's test of sphericity and the KMO Measure of Sampling Adequacy are very similar, but the latter provides various ranges of acceptable variance rather than simply significant or nonsignificant. Generally, the metrics are as follows:

- 0.00 through 0.50 – Unacceptable – Bad
- 0.50 through 0.60 – Miserable – Bad
- 0.60 through 0.70 – Mediocre – Okay
- 0.70 through 0.80 – Middling – Okay
- 0.80 through 0.90 – Meritorious – Good
- 0.90 through 1.00 – Marvelous – Great

Given these, authors should seek KMO Measure of Sampling Adequacy above .60 before performing their EFA (Dziuban & Shirkey, 1974; Kaiser, 1970). If a lower value is obtained, variables with small intercorrelations can be removed to improve suitability for EFA.

In addition to these two methods, authors should ensure that their sample size is sufficient, and scholars have suggested an array of guidelines. Among the most popular recommendations for minimum sample size are 200 to 500 participants (depending on communalities and other factors; Comrey & Lee, 1992; MacCallum, Widaman, Zhang, & Hong, 1999) and between a 5-to-20 and participant-to-variable ratio (Costello & Osborne, 2005; Hair, Black, Babin, Anderson, & Tatham, 2006; Velicer & Fava, 1998). Of these, it appears that most authors are accepting of a minimum sample size of 200 and 5-to-1 participant-to-variable ratio, whichever is greater. Although some may consider this cutoff conservative, the current article supports its use for most EFA applications.

1.2. Factor Analytic Method

Next, authors must choose their preferred factor analytic method, which is likely the most contested aspect of EFA. Multiple methods exist, and among the most common are Principal Components Analysis (PCA), Principal Axis Factoring (PAF), and Maximum Likelihood (ML). Each of these methods demonstrates particular strengths and weaknesses.

PCA is often considered the most popular "EFA" method; however, authors have repeatedly speculated that its popularity is not due to its statistical properties but rather because it is the default factor extraction methods in SPSS (Costello & Osborne, 2005). Further yet, authors have consistently condemned the use of PCA for several reasons (Fabrigar et al., 1999; Preacher & MacCallum, 2003). Primarily, PCA is not a true form of factor analysis. Instead, it is based upon a different mathematical model than EFA techniques, and disparate results arise due to multiple model differences. First, PCA does not account for the structure of correlations but only attempts to explain as much variance as possible within the measured variables (Fabrigar & Wegener, 2012). For this reason, PCA does not provide results on latent constructs but only serves to account for the variance of measured variables (Fabrigar et al., 1999; Preacher & MacCallum, 2003). Second, EFA methods account for common variance and unique variance within measured variables, and resultant factors are derived from common variance alone. Alternatively, PCA does not make a distinction between these two forms of variance, and components represent common and unique variance. Some authors propose that PCA may be analogous to an EFA method, which assumes that measured variables' unique variances are zero. In the same breath, however, these authors also note that these assumptions are often unrealistic, as virtually all measured variables will include measurement error (Bentler & Kano, 1990; Fabrigar & Wegener, 2012). With these aspects taken into consideration, authors should not apply PCA when performing EFA, as the results do not reflect true factors.

Alternatively, many authors suggest PAF for EFA. The goal of PAF is to produce a set of factor loading estimates that comes as close as possible to reproducing the common variance within a correlation matrix (De Winter & Dodou, 2012). Several authors have detailed the mathematics that provide these estimates (Cudeck, 2000; Gorsuch, 1988). These mathematics are not discussed in the current article, as they are automated within modern statistical programs. Nevertheless, authors should be aware of certain aspects of PAF. First, modern PAF methods must undergo several statistical iterations before arriving at a final solution. Second, once the final solution has been obtained, a number of alternative solutions may reproduce the common variance equally well, and the initially derived factor loadings are chosen for computing ease rather than interpretability (Fabrigar & Wegener, 2012). For this reason, authors should subsequently rotate their PAF solutions to improve interpretability while retaining correctness. Third, PAF assumes that unique variances (errors) are normal but does not assume that variables are multivariate normal with linear interrelationships (Costello & Osborne, 2005; De Winter & Dodou, 2012). Together, these factors prompt PAF to provide accurate results in most situations.

Last, ML is another suggested EFA possibility. ML mathematically determines the factor loading and unique variance estimates that are most likely to have produced the observed data (De Winter & Dodou, 2012). Once again, these mathematics are not described within the current article, but they have been detailed elsewhere (Cudeck, 2000; Gorsuch, 1988). Six aspects should be known about ML. First, ML must also undergo several statistical iterations before arriving at a final solution. Second, the final solution must be rotated for interpretability. Third, ML assumes that unique variances (errors) are normal. Fourth, ML assumes that variables are multivariate normal with linear interrelationships, which is not an assumption of PCA or PAF (Costello & Osborne, 2005; De Winter & Dodou, 2012). Fifth, ML is more likely to produce inaccurate or improper solutions than PAF, but these cases are generally considered rare (Fabrigar & Wegener, 2012). Sixth, ML can provide an array of model fit indices which are shared with CFA/SEM. Unlike other EFA methods, ML model fit indices allow direct comparisons between multiple EFA solutions and enables hypothesis testing (Conway & Huffcutt, 2003; Furr, 2011). Overall, ML can provide more information than other EFA methods, but it also includes stricter assumptions.

Overall, three recommendations can be given for factor analysis method choices. First, authors should not apply PCA for EFA. Second, the added information provided by ML is useful, but its drawbacks may result in inaccurate or inappropriate results. For this reason, authors should perform PAF if they *are not* interested in model fit indices. Alternatively, they should perform ML if they *are* interested in model fit indices; however, authors should compare their ML results against PAF results to ensure that the two are similar and ML assumptions are not violated, resulting in inaccurate or inappropriate inferences. Third, other factor analytic methods exist, but they are less common and possess statistical assumptions of their own. No matter the method applied, authors should be knowledge of their benefits and limitations.

1.3. Factor Retention Method

The goal of EFA is to produce a number of factors that accurately and understandably explain the observed correlation matrix. The resultant number of factors should not be benefitted by adding another factor, and the model should perform substantially worse if a factor is removed. For this reason, authors need to be extremely careful in determining their factor retention decisions, and an array of factor retention methods exist. These include the Kaiser criterion, visual scree plot (VSP) analysis, parallel analysis, and Velicer's Minimum Average Partial (MAP) test. Most of these methods create factor cutoffs through analyzing eigenvalues, which are numerical values representing the "variances in measured variables accounted for by each of the common factors" (Fabrigar & Wegener, 2012, p. 45). Eigenvalues are calculated by summing the squared factor loadings (Kline, 2014). For this reason, factors with small eigenvalues represent little common variances and should not be included in analyses.

Possibly the first factor retention method is the Kaiser (1960) criterion, which suggests that all factors with eigenvalues above 1 should be retained. The justification for the "eigenvalues-greater-than-one" rule was initially derived from the mathematics behind PCA (Cliff, 1988; Yeomans & Golder, 1982), prompting concerns regarding its use within EFA. In addition to theoretical doubts, the Kaiser criterion also demonstrates practical concerns. Particularly, factors with almost identical eigenvalues, such as 1.01 and 0.99, receive different retention decisions, despite explaining almost equal amounts of common variance. These concerns are well known, and authors have repeatedly condemned the Kaiser criterion. Costello and Osborne (2005) noted that it is "amongst the least accurate methods" for factor retention decisions (p. 2), which is supported by a multitude of simulation studies (Kline, 2014; Patil, Singh, Mishra, & Donavan, 2008). For this reason, authors regularly apply other factor retention methods.

One of the most popular factor retention methods is the VSP analysis (Zoski & Jurs, 1996). To perform a VSP analysis, authors plot each eigenvalue on a graph and determine when decreases in successive eigenvalues become less noticeable—called an "elbow." Each factor with an eigenvalue before the elbow is retained, as these factors are considered to represent common variance significantly better than factors after the elbow. The primary benefit of a VSP analysis is also its drawback; it is easy to perform and does not require any statistical analyses. Due to this, VSP analyses are greatly subject to interpretation and may not provide clear cutoffs in certain situations. Nevertheless, due to its accuracy beyond the Kaiser criterion and the need for simple cutoffs, psychometricians are generally accepting of VSP analyses.

Another method is parallel analysis. Parallel analysis involves the generation of many randomized data sets with the same number of variables and cases as the data set of interest (Hayton, Allen, & Scarpello, 2004). Then the randomized data sets are subjected to EFA, and the (a) average or (b) the 95th percentile of each eigenvalue is recorded (depending on the author's choice). Once this process has been repeated a specified number of times, the recorded eigenvalues are compared to the eigenvalues derived from the EFA performed upon the data set of interest. Each factor with an eigenvalue greater than its respective pair obtained from the randomized data sets is retained. Originally, parallel analyses were difficult to complete, but scholars have provided easy-to-use programs to automate the process. In fact, Patil, Singh, Mishra, and Donavan (2007) even developed an extremely simple web applet that can be used without statistical software programs, providing access to virtually anyone.

Further, the logic of a parallel analysis is simple. When performing an EFA upon a randomized data set, the resulting eigenvalues cannot represent meaningful factors. Instead, they represent statistical artifacts (Horn, 1965; O'Connor, 2000). Any factor with an eigenvalue greater than its respective random-data eigenvalue is believed to provide explanatory information beyond statistical artifacts alone, whereas a factor with an eigenvalue less than its respective random-data eigenvalue cannot be differentiated from statistical artifacts alone.

Although parallel analysis is theoretically sound, it incurs particular practical concerns. Like the Kaiser criterion, disparate retention decisions can be obtained for almost identical factor eigenvalues. For example, a parallel analysis may indicate that the third-factor eigenvalue must exceed 1.10 to be retained. Although 1.11 would meet this cutoff and 1.09 would not, it is debatable whether these values actually demonstrate a significant difference. Despite these concerns, simulations have generally supported parallel analysis for factor retention decisions (Hayton et al., 2004).

Last, Velicer's MAP is another often-used factor retention method (O'Connor, 2000; Zwick & Velicer, 1986). Velicer's MAP begins by calculating the average squared correlation of the original data set. Then the variance for each factor is successively partialed out, and the averaged squared correlation of each revision is calculated. The step that resulted in the lowest averaged squared correlations determines the number of factors to retain, as this step removed or explained the most amount of variance within the intercorrelations. Multiple studies have demonstrated that Velicer's MAP is accurate in determining factor retention decisions, but it sometimes has the tendency to underestimate the number of factors (O'Connor, 2000; Zwick & Velicer, 1986). Nevertheless, the method still appears to be an appropriate guide for factor retention decisions.

Overall, it is recommended that authors no longer apply the Kaiser criterion. Instead, authors should use VSP analysis in conjunction with parallel analysis, Velicer's MAP, or both. This suggestion is very reasonable given the current accessibility of these methods.

1.4. Factor Rotation Method

Once the number of factors has been chosen, the individual variable loadings need to be interpreted; however, the initial results are difficult to analyze. As mentioned, many EFA methods produce a valid factor matrix that is simple to compute, but many other factor matrices may also be correct and easier to interpret (Fabrigar & Wegener, 2012). For this reason, authors must rotate their EFA solutions. Several methods exist, and these fall into two categories: orthogonal and oblique. Orthogonal rotations do not allow the resultant rotated factors to be correlated, which is often *not* preferable (Costello & Osborne, 2005; Fabrigar et al., 1999; Hinkin, 1995, 1998). Many scales are multidimensional. It is likely that these subdimensions are related, and EFA results should be able to account for these interrelations. Alternatively, oblique rotations allow for the resultant factors to be correlated, causing authors to be preferable toward oblique rotations (Costello & Osborne, 2005; Fabrigar et al., 1999; Hinkin, 1995, 1998). Forcing related factors to be uncorrelated (i.e. orthogonal rotations) is seen as more harmful than allowing, but not forcing, unrelated factors to correlate (i.e., oblique rotations; Fabrigar & Wegener, 2012). Nevertheless, scholars should be knowledgeable of both types of rotation.

Orthogonal rotation methods include quartimax and varimax. The first created rotation method was quartimax, but authors quickly discovered its several practical drawbacks (Kaiser, 1958). For this reason, Kaiser (1958) created varimax rotation. Varimax rotation seeks to increase the variances of the factor loadings, resulting in both large and small factor loadings. This is often preferable, as variables will (hopefully) clearly load or not load onto each factor. Since its inception, Varimax has continues to be a popular and useful rotation method.

On the other hand, oblique rotation methods include promax and direct oblimin. Promax was among the first created oblique rotation methods (Cureton & Mulaik, 1975). A promax rotation begins by performing a varimax rotation, and then it allows the factors to correlate through raising the factor loadings to a specified power. A value of 4 is most often used, likely due to the default promax procedure in SPSS. Promax, along with several other oblique rotations, are considered indirect methods of rotations, as they alter the results of other rotation methods (Fabrigar & Wegener, 2012). Alternatively, direct oblimin rotations directly rotate to their final solution (Fabrigar & Wegener, 2012). Although several authors believe that direct oblimin is a rotation method, the term actually refers to a family of rotations. Each direct oblimin rotation is defined by its delta value, which specifies the extent that the factors may correlate (Browne, 2001; Jennrich & Sampson, 1966). For instance, a delta of zero places equal weighting of correlated and uncorrelated factors, and it is called a direct quartimin rotation. Direct quartimin is likely the most popular direct oblimin rotation method, as it is the default direct oblimin rotation in SPSS, and it is probably the rotation method that authors intend when they mention direct oblimin. Authors should be wary of applying other delta values when performing direct oblimin rotations, as not all of these rotations perform equally as well. Further, from several simulation studies, authors have relatedly demonstrated the satisfactory performance of direct quartimin (Browne, 2001; Ford, MacCallum, & Tail, 1986; Hinkin, 1995, 1998).

Together, the rotation choice should be guided by a priori theory. If an author believes that the resultant factors are uncorrelated, they should perform an orthogonal rotation. Currently, varimax seems to be the most preferred. Alternatively, if an author believes that the resultant factors are correlated, then they should perform an oblique rotation. A direct oblimin rotation with a delta of zero, also called a direct quartimin, seems to be the most preferred.

1.5. Factor Loading Cutoff

In addition to determining the number of factors within a set of variables, factor analysis also determines the extent that each variable represents each emergent factor through loading values. Often, variables representative of a factor are retained, whereas those representative of multiple factors or no factors at all are removed (Hair et al., 2006; Hinkin, 1995, 1998; Tabachnick & Fidell, 2007). Several questions arise in interpreting the factor loadings to make these decisions. First, it is unclear when a variable sufficiently loads onto a factor and is considered representative of that factor. Second, it is also unclear when a variable loads onto too many factors and is considered not clearly representative of any. Fortunately, authors have provided several cutoffs for these decisions.

Possibly the most popular cutoff for "good" factor loadings onto a primary factor is 0.40 (Hinkin, 1995, 1998), but other authors have proposed values of 0.30 (Costello & Osborne, 2005), 0.32 (Tabachnick & Fidell, 2001), 0.45 (Tabachnick & Fidell, 2007), and beyond (Hair et al., 2006). Alternatively, the criterion for "excessive" factor loadings onto alternative factors is even less decisive. Authors have proposed that alternative factor loadings should not exceed a certain value, such as 0.32 or .40 (Costello & Osborne, 2005). Other authors have proposed that differences between the primary and alternative factor loadings should be considered, such as a loading difference of 0.20 between the primary and alternative factors (Hinkin, 1998). The current article suggests a combination of these proposals. It is recommended that satisfactory variables (a) load onto their primary factor above 0.40, (b) load onto alternative factors below 0.30, and (c) demonstrate a difference of 0.20 between their primary and alternative factor loadings. The current article dubs this the .40–.30–.20 rule.

2. CURRENT PRACTICES

With the primary decisions involved in EFA reviewed, the current study provides an overview of the current practices of EFA within cyberpsychology and human–computer interaction studies. To perform this overview, every article that mentioned "factor analysis" in seven journals was retrieved, and their EFA decisions were recorded by the primary author. The seven journals were *Cyberpsychology, Behavior, and Social Networking; Computers in Human Behaviors; Journal of Educational Computing Research; Journal of Computer-Mediated Communication; New Media and Society; International Journal of Human–Computer Interaction*; and *Computers & Education*. From this search, 220 articles were discovered, and 163 reported an EFA.

It should be noted that, if multiple EFAs were performed within the same article, the decisions were recorded only once. This was to prevent articles that performed several EFAs from being overrepresented and skewing the results; however, exceptions were made for two occurrences. First, if the authors chose different decisions for the factor analyses, the multiple decisions were recorded. Second, if the authors used multiple samples but the same decisions, the decisions were recorded only once, but each separate sample size was recorded.

2.1. Data Inspection Techniques

The application of two data inspection techniques, KMO Measure of Sampling Adequacy and Bartlett's test of sphericity, was recorded. Also, the overall sample size was recorded. Participant-to-item ratios were meant to be included, but many studies were unclear about how many variables were included within their factor analyses. For this reason, participant-to-item ratios were not recorded.

2.2. Factor Analytic Method

The chosen factor analytic method was recorded. The recorded options were PCA, PAF, ML, other, and unspecified.

2.3. Factor Retention Method

The factor retention method was recorded. The recorded options were the Kaiser criterion, VSP analysis, parallel analysis, Velicer's MAP, factor variances, number of representative items, other, and unspecified.

2.4. Factor Rotation Method

The factor rotation method was recorded. First, it was determined whether the method was orthogonal, oblique, or unspecified. Second, it was noted whether the method was varimax, quartimax, equimax, direct oblimin, promax, unrotated, or unspecified.

2.5. Factor Loading Cutoff

The factor loading cutoff was recorded. First, it was recorded whether the author stated a cutoff for the loading upon the primary factor and the specified value itself was recorded. Next, if the author did not specify the primary factor loading cutoff, the lowest primary factor loading of all included variables was recorded. This value was considered the inferred factor loading cutoff.

Second, it was recorded whether the author stated a cutoff for the loading upon alternative factors. Then, the stated cutoff value was recorded, separated by whether the cutoff was a stand-alone value or distance from the primary factor loading. If the author did not mention a cross-loading cutoff, no inferred value was recorded. This is because most authors do not list alternative factor loadings within their tables, and the vast majority of alternative value cutoffs could not be determined.

3. RESULTS

Authors' choices for each EFA decision within 163 articles were recorded. Given that statistical standards and practices change fairly quickly, the results reported within tables are separated by all articles and those published within the past decade (starting in 2005). Of articles published in 2005 or later, 130 included an EFA. In next describing the results, only those published within the last decade are referenced.

Table 1 provides an overview of data inspection techniques and sample sizes. Overall, KMO Measure of Sample Adequacy (42%) and Bartlett's Test of Sphericity (39%) were applied almost equally. Although not evident in the table, almost every author who applied one method also applied the other (90%). Despite this, most authors did not apply either method (56%).

In addition, the mean and median sample sizes were 402 and 253, respectively. Of these, the median should be considered, as the mean is greatly skewed by a few large sample sizes. Also, the standard deviation of 568 indicates that the sample sizes varied greatly across studies. This is also reflected in the minimum and maximum values of 31 and 5,509. To gain a better understanding of sample size choices, a histogram of sample sizes is included in Figure 1, lists the number of articles that fall into certain sample size ranges. It seems that many articles did not meet the recommended number of 200 participants (37%).

Table 2 provides an overview of factor analytic method decisions. The most popular method was PCA (53%), although it is not an EFA method, followed by PAF (20%). Beyond these two, authors occasionally performed ML (12%) or another method (3%). Surprisingly, a large number of authors did not report their

TABLE 1
Chosen Data Inspection Techniques and Sample Sizes

	All Articles	Within Past Decade
1. KMO Measure of Sample Adequacy	58 (36%)	54 (42%)
2. Bartlett's Test of Sphericity	52 (32%)	51 (39%)
1. *M* sample size	379	402
2. *Mdn* sample size	244	253
3. Sample size *SD*	524	568
4. Minimum sample size	20	31
5. Maximum sample size	5,509	5,509
1. No. of participants between 1 and 100	25 (15%)	16 (12%)
2. No. of participants between 101 and 200	40 (24%)	33 (25%)
3. No. of participants between 201 and 300	36 (21%)	29 (22%)
4. No. of participants between 301 and 400	18 (11%)	14 (10%)
5. No. of participants between 401 and 500	15 (09%)	13 (10%)
6. No. of participants above 501	34 (20%)	29 (22%)

Note. The total number of reviewed articles was 163. The total number of included sample sizes was 168. The total number of reviewed articles within the past decade was 130. The total number of included sample sizes within the past decade was 134. The sum of each column may be greater than the total number of articles, as a single article may contain multiple samples and chosen decision choices. KMO = Kaiser–Meyer–Olkin.

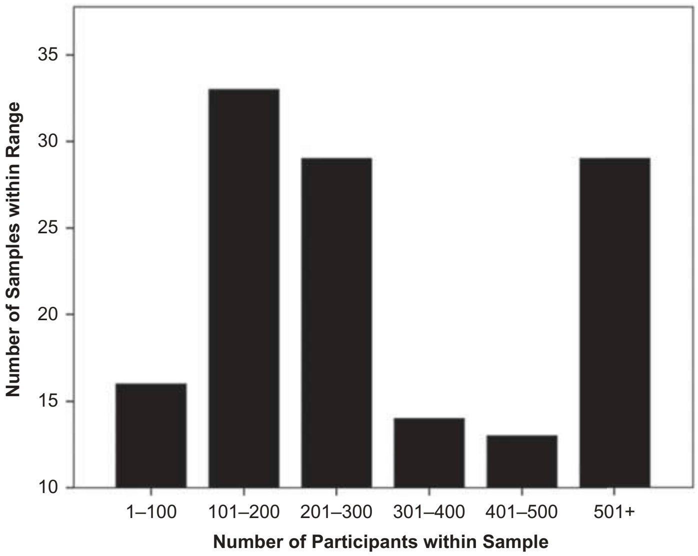

FIG. 1. Histogram of sample size choices.

TABLE 2
Chosen Factor Analytic Methods

	All Articles	Within Past Decade
1. Principal components analysis	85 (52%)	69 (53%)
2. Principal axis factoring	33 (20%)	26 (20%)
3. Maximum likelihood	15 (9%)	15 (12%)
4. Other method	4 (2%)	4 (3%)
5. Did not specify	34 (21%)	23 (18%)

Note. Total number of reviewed articles was 163. The total number of reviewed articles within the past decade was 130. The sum of each column may be greater than the total number of articles, as a single article may contain multiple samples and chosen decision choices.

factor analytic method (18%). In most of these cases, the author simply stated that they performed "factor analysis."

Table 3 provides an overview of factor retention decisions. Of these, the Kasier criterion was the most popular method (54%). VSP analysis (33%) and parallel analysis (10%) were also regularly used. Velicer's MAP was only sporadically applied (4%). Two undiscussed methods, amount of variance explained (5%) and number of representative items (5%), were also occasionally used as factor retention decisions. Last, authors sometimes used other methods (3%), and many did not specify their method at all (32%). It should be noted that, of those that reported their method, 45% of authors used multiple methods, and 42 percent of those which applied the Kaiser criterion also applied VSP analysis, parallel analysis, or Velicer's MAP.

TABLE 3
Chosen Factor Retention Method

	All Articles	Within Past Decade
1. Kaiser Criterion	84 (52%)	70 (54%)
2. VSP analysis	50 (31%)	43 (33%)
3. Parallel analysis	13 (8%)	13 (10%)
4. Velicer's MAP	5 (3%)	5 (4%)
5. Variance explained	8 (5%)	7 (5%)
6. No. of representative items	8 (5%)	6 (5%)
7. Other	4 (2%)	4 (3%)
8. Did not specify	57 (35%)	41 (32%)

Note. Total number of reviewed articles was 163. The total number of reviewed articles within the past decade was 130. The sum of each column may be greater than the total number of articles, as a single article may contain multiple samples and chosen decision choices. VSP = visual scree plot; MAP = Minimum Average Partial.

TABLE 4
Chosen Factor Rotation Method

	All Articles	Within Past Decade
1. Orthogonal	94 (58%)	76 (58%)
1a. Varimax	88 (94%)	70 (92%)
1b. Quartimax	1 (1%)	1 (1%)
1c. Equimax	1 (1%)	1 (1%)
2. Oblique	52 (32%)	42 (32%)
2a. Direct Oblimin	27 (52%)	21 (50%)
2b. Promax	12 (23%)	11 (26%)
3. Unrotated	2 (4%)	0 (0%)
4. Did not specify	25 (15%)	19 (15%)

Note. Total number of reviewed articles was 163. The total number of reviewed articles within the past decade was 130. The sum of each column may be greater than the total number of articles, as a single article may contain multiple samples and chosen decision choices.

Table 4 provides an overview of factor rotation methods. Orthogonal rotations (58%) were more popular than Oblique rotations (32%). Of the orthogonal rotations, varimax is the most popular (92%). Of the oblique rotations, direct oblimin rotations (50%) were the most popular followed by promax (26%). Many did not specify their rotation method (15%).

Last, Table 5 provides an overview of the chosen factor loading cutoffs. Of those that directly stated a cutoff, the average value was 0.44. A surprisingly large number of authors did not specify their cutoff for the primary factor loading (46%). Of those that specified their cutoff or could be inferred, a large variation was seen in values, as seen in Table 5. The average was high and appropriate at 0.48, but a portion of studies used cutoffs 0.30 or lower (10%). Further, most authors made no mention of cross-loading cutoffs (78%). Of those that did, they

TABLE 5
Chosen Factor Loading Cutoff

	All Articles	Within Past Decade
1. Stated primary loading were cutoff	83 (51%)	71 (54%)
2. Did not specify	81 (49%)	60 (46%)
3. Stated primary loading cutoff value mean	.43	.44
4. Inferred primary loading cutoff value mean	.47	.48
5. Primary loading cutoff between 0.00 and 0.30[a]	20 (12%)	13 (10%)
6. Primary loading cutoff between 0.301 and 0.40[a]	49 (30%)	37 (28%)
7. Primary loading cutoff between 0.401 and 0.50[a]	42 (26%)	39 (30%)
8. Primary loading cutoff between 0.501 and 0.60[a]	26 (16%)	24 (18%)
9. Primary loading cutoff above 0.601[a]	5 (3%)	4 (3%)
1. Stated cross-loadings were cutoff	33 (20%)	29 (22%)
2. Did not specify	130 (80%)	101 (78%)
3. Cross-loading cutoff values	.20–5 (3%)	.20–5 (4%)
	.30–2 (1%)	.30–2 (2%)
	.35–1 (1%)	.35–1 (1%)
	.45–1 (1%)	.45–1 (1%)
	.50–2 (1%)	.50–2 (2%)
4. Cross-loading cutoff difference values	.10–3 (2%)	.10–3 (2%)
	.30–1 (1%)	.30–1 (1%)

Note. Total number of reviewed articles was 163. The total number of recorded primary factor cutoffs was 164, as one article made different factor cutoff decisions between its first and second sample. The total number of reviewed articles within the past decade was 130. The total number of recorded primary factor cutoffs was 131, as one article made different factor cutoff decisions between its first and second sample. The sum of each column may be greater than the total number of articles, as a single article may contain multiple samples and chosen decision choices.

[a]Values include stated and inferred cutoffs.

most often specified that they analyzed cross-loadings, but never made a mention of cutoff values (52%). Of those that did mention a value, most chose a certain cutoff rather than a distance value (73%). The most common cutoff was a loading of 0.20. The most common distance value was 0.10.

4. DISCUSSION

Creating psychometrically sound and valid scales is particularly important for cyberpsychology and human–computer interaction scholarship. Both of these fields are interested in latent constructs, and few established measures exist due to their relative newness. For this reason, the current article has two primary sections. First, a review provided suggestions for five important statistical and methodological decisions during the EFA process. Second, an overview detailed authors' EFA decisions to determine whether appropriate methods are being used. From these two sections, several inferences can be made.

First, most authors do not investigate their data for violations of EFA assumptions. This is surprising, as KMO Measure of Sample Adequacy and Bartlett's Test of Sphericity have existed for decades (Bartlett, 1950; Dziuban & Shirkey, 1974; Kaiser, 1970) and may denote when a set of variables may demonstrate small correlations that are unsuitable for EFA. It is highly suggested that future authors incorporate these two data investigation techniques into research and practice, as it may prevent inappropriate applications of EFA.

Second, the mean sample size of articles is not overly concerning (402 participants); however, this value is skewed by several large outliers, and other values should be considered. When analyzing the median sample size (253 participants), it is below many recommended cutoffs (Comrey & Lee, 1992; MacCallum et al., 1999). In regards to the cutoff suggested by the current article, 200 participants, many articles failed to meet this recommendation (37%). Further, the lowest observed sample size was 31—much lower than the recommended cutoff of any author (Hair et al., 2006; Velicer & Fava, 1998). Overall, authors should strongly consider their sample sizes, as EFA may be greatly inaccurate with an insufficient number of participants.

Third, the most common factor analytic method was PCA, which is not even a true form of EFA and presents several statistical concerns (Fabrigar et al., 1999; Preacher & MacCallum, 2003). Alternatively, PAF and ML were much less common, despite their statistical advantages (Costello & Osborne, 2005; De Winter & Dodou, 2012). Authors should become familiar with these latter two methods, as their properties can greatly benefit research and provide more accurate study results. Also, a surprising number of authors did not even mention their factor analytic method, and only stated that they performed "factor analysis." This is extremely problematic, as readers cannot be certain of the validity of results and replication becomes impossible. Future researcher should always note their chosen factor analytic method.

Fourth, the Kaiser criterion is the most popular factor retention method, despite a long history of voiced concerns (Kline, 2014; Patil et al., 2008). Even more problematic, most authors who applied the Kaiser criterion did not perform any other factor retention methods to support their results. Within these studies, authors' chosen number of retained factors may be grossly overestimated. Also, like factor analytic method decisions, a large number of authors did not specify their factor retention methods—almost large as the number of authors who applied VSP analysis. Once again, this is extremely problematic for the same reasons just noted. It is highly suggested that authors begin applying and reporting more statistically sound factor retention methods, such as VSP analysis, parallel analysis, and Velicer's MAP.

Fifth, surprisingly, orthogonal rotations were more popular than oblique rotations. This is not a problem itself, per se, but most authors did not specify the theoretical justification for their rotation method. For this reason, it appears that authors may be unsure of the differing dynamics in the two categories of rotations and the statistical concerns that may arise from performing the incorrect type of rotation. Although it may not be a sound practice, it is suggested that authors perform oblique rotations when they are unsure toward the correct rotation method, as forcing factors to be uncorrelated may be more problematic than allowing them to correlate (Fabrigar & Wegener, 2012). Further, of the orthogonal rotations, varimax was the most popular. Of the oblique rotations, direct oblimin rotations were the most popular, but a considerable number of authors also applied a promax rotation. Once again, a large number of authors did not specify their rotation or simply stated the category of rotation (i.e., orthogonal or oblique rotation). This is a poor practice and should be rectified.

Sixth, the number of authors who did and did not state their primary factor loading cutoff was almost even. Of those who did state their cutoff, the average value was 0.44. This value is in accordance to priori suggestions and appropriate. Many articles, although not directly stated, could have their primary factor loading cutoff inferred from included tables. The average value of inferred cutoffs was 0.48. It is unclear why these authors would not state their cutoff, as most met or exceeded prior suggestions. Further, when analyzing stated and inferred cutoffs together, almost half chose a value above 0.40. It seems that authors' primary factor loading cutoffs are generally appropriate and not a large area of concern.

Seventh, the large majority of authors made no mention of variable cross-loadings. Of those who mentioned cross-loadings, about half only stated that variables could not cross-load but did not provide a cutoff value. Ensuring that variables do not include alternative constructs of interest is extremely important for research and practice. If an item, for instance, is included within a scale but cross-loads with another scale, the observed relationship between these two scales would be artificially inflated due to the criterion contamination of the item(s). Of all decisions made within the EFA process, it seems that variable cross-loadings were the least considered by authors. Once again, it is strongly suggested that future researchers directly state their cross-loading cutoff values for variable retention, and the suggested .40–.30–.20 rule could be easily applied. This rule recommends that satisfactory variables (a) load onto their primary factor above 0.40, (b) load onto alternative factors below 0.30, and (c) demonstrate a difference of 0.20 between their primary and alternative factor loadings.

Eighth, it seems that standards and practices have not greatly changed within the past decade when performing EFA. Through comparing all articles and only those published within the past decade in Tables 1 through 5, few notable changes can be seen. This is surprising, as computing technology has improved and statistical software has become more user-friendly. Authors within the study of cyberpsychology and human–computer interaction should take advantage of these technological improvements, as they should be the most adaptable to new technology of any research field.

Taken together, several systematic issues are present in studies performing EFA. To overcome these issues, the current article provides several suggested best practices. Table 6 is included to summarize these recommended best practices and provide a checklist to authors. It is recommended that researchers performing EFAs should follow this checklist, step by step, to ensure success in performing and reporting their methods. In addition to these recommend practices, future considerations should be noted.

5. DIRECTIONS FOR FUTURE RESEARCH

The current study has only two, albeit broad, suggestions for future research. Authors should continue to denote the importance of adequate measurement within research (Bunz, 2004; Faiola, Ho, Tarrant, & MacDorman, 2011; Yang, Linder, & Bolchini, 2012). Particularly, future research could provide suggestions toward several other aspects of the scale creation and evaluation process, such as item writing or confirmatory factor analysis. Also, authors could also provide similar overviews of alternative attributes of scales. For instance, the current article noted very few suggestions toward the state of validity inferences within research. It may be fruitful to review current studies and determine the methods by which authors ensure the validity of their measures. Then suggestions could be provided to enhance the theoretical basis of extant scales.

In addition, a great need is still evident for psychometrically sound and valid measures within the study of cyberpsychology and human–computer interaction. As Howard and Jayne (2015) demonstrated, a great number of research results are based upon measures with limited investigations

TABLE 6
Checklist of Suggested Recommendations for EFA

EFA Decision	Recommendation	Checklist
1. Data inspection	Is your sample size greater than 200 participants and a 5-to-1 participant-to-variable ratio?	☐
	Did you perform a Bartlett's Test of Sphericity and/or KMO Measure of Sampling Adequacy?	☐
2. Factor analytic method	Did you avoid using principal components analysis?	☐
	Did you perform the factor analytic method with the fewest statistical assumptions while still providing all desired information?	☐
3. Factor retention method	Did you avoid using the Kaiser criterion (Eigenvalue $<$ 1 rule)?	☐
	Did you perform a visual scree plot analysis *and* parallel analysis, Velicer's MAP, or both?	☐
4. Factor rotation method	Did you use a rotation method which aligns with the expected correlation of your factors?	☐
	If using an orthogonal rotation, did you perform a varimax rotation? Alternatively, if using an oblique rotation, did you perform a direct quartimin rotation?	☐
5. Factor loading cutoff	Do your items load more than .40 on their primary factor?	☐
	Do your items load below .30 on their alternative factors?	☐
	Do your items have a greater difference than .20 between their primary and alternative factor loadings?	☐

Note. KMO = Kaiser–Meyer–Olkin; MAP = Minimum Average Partial.

into their psychometric properties and validity, and these study results may be inaccurate due to the use of these measures. Although the current article may seem condemning of current scale development and evaluation research, authors should certainly persist in their efforts. Indeed, authors should apply the recommendations provided within the current article, and create a new wave of methodologically sound studies. Through this effort, multiple areas of research could be greatly improved and studies' inferences could be more reliable.

Currently, some areas of research have already seen interest in developing reliable and valid scales. When reviewing EFA practices in the current article, among the most studied constructs were Internet addiction (Caplan, 2002; Pawlikowski, Altstötter-Gleich, & Brand, 2013), cyberbulling (Çetin, Yaman, & Peker, 2011; Lam & Li, 2013), e-learning perceptions (Ozkan & Koseler, 2009; van Braak & Tearle, 2007), and attitudes toward technology (Korobili, Togia, & Malliari, 2010; Rainer & Miller, 1996). Researchers disagree about the dimensionality and conceptualization of these constructs, partially due to the differing EFA decisions across articles that can greatly impact outcomes and interpretation. Future research using the recommended best practices of the current article may clarify the dimensionality and conceptualization of these constructs. In addition, computer self-efficacy (Howard, 2014), blogging motivations (Baker & Moore, 2011), technology-related emotions (Charlton & Birkett, 1995), unethical computer behaviors (Namlu & Odabasi, 2007), and presence (Qin, Rau, & Salvendy, 2009) were also studied but to a lesser extent. In future scale creation efforts, it may be beneficial to reexamine these constructs to provide clarity in ongoing research debates. Nevertheless, value is also present in analyzing measurement issues for lesser researched constructs, as creating a novel and useful scale may prompt further research.

6. CONCLUSION

The current article provided an overview of five primary decisions made when performing an EFA and reviewed all EFA decisions of authors within seven cyberpsychology or human–computer interaction journals. The results demonstrated that multiple systematic and problematic concerns exist in regards to these EFA decisions. Particularly, many authors do not check their data for violations of EFA assumptions, sample sizes are routinely below suggested cutoffs, PCA is the most popular factor analytic method, the Kaiser criterion is the most popular factor retention method, justifications are sparsely noted for rotation decisions, and variable cross-loadings are rarely considered. Possibly the only consistently favorable aspect of authors' EFA decisions is their chosen primary factor loading cutoff. For these reasons, it is strongly recommended that authors take note of the suggestions provided within the current article toward EFA decisions. Also, future research should provide other suggestions toward scale creation and evaluation, and authors should remain attentive toward creating psychometrically sound and valid measures.

ACKNOWLEDGMENTS

I thank Laura Vrana for her comments on a previous version of this article.

REFERENCES

Baker, J. R., & Moore, S. M. (2011). Creation and validation of the personal blogging style scale. *Cyberpsychology, Behavior, and Social Networking*, *14*, 379–385.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*, 77–85.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* (Methodological), 289–300.

Bentler, P. M., & Kano, Y. (1990). On the equivalence of factors and components. *Multivariate Behavioral Research*, *25*, 67–74.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*, 111–150.

Bunz, U. (2004). The computer-email-web (CEW) fluency scale-development and validation. *International Journal of Human–Computer Interaction*, *17*, 479–506.

Caplan, S. E. (2002). Problematic Internet use and psychosocial well-being: Development of a theory-based cognitive–behavioral measurement instrument. *Computers in Human Behavior*, *18*, 553–575.

Çetin, B., Yaman, E., & Peker, A. (2011). Cyber victim and bullying scale: A study of validity and reliability. *Computers & Education*, *57*, 2261–2271.

Charlton, J. P., & Birkett, P. E. (1995). The development and validation of the Computer Apathy and Anxiety Scale. *Journal of Educational Computing Research*, *13*, 41–59.

Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin*, *103*, 276–279.

Comrey, A. L. & Lee H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6*, 147–168.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*, 1–9.

Cudeck, R. (2000). Exploratory factor analysis. *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, pp. 265–296.

Cureton, E. E., & Mulaik, S. A. (1975). The weighted varimax rotation and the promax rotation. *Psychometrika*, *40*, 183–195.

De Winter, J. C., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*, 695–710.

Dziuban, C. D., & Harris, C. W. (1973). On the extraction of components and the applicability of the factor model. *American Educational Research Journal*, *10*, 93–99.

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, *81*, 358–361.

Fabrigar, L. R., & Wegener, D. T. (2012). *Understanding statistics: Exploratory factor analysis*. New York, NY: Oxford University.

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.

Faiola, A., Ho, C. C., Tarrant, M. D., & MacDorman, K. F. (2011). The aesthetic dimensions of US and South Korean responses to web home pages: A cross-cultural comparison. *International Journal of Human–Computer Interaction*, *27*, 131–150.

Ford, J. K., MacCallum, R. C., & Tail, M. (1986). The applications of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*, 291–314.

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. Thousand Oaks, CA: Sage.

Gorsuch, R. L. (1988). Exploratory factor analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 231–258). New York, NY: Springer.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, *68*, 532–560.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis (Vol. 6)*. Upper Saddle River, NJ: Pearson Prentice Hall.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale–2. *Computers in Human Behavior*, *40*, 133–143.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*, 191–205.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*, 393–416.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, *21*, 967–988.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*, 104–121.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185.

Howard, M. C. (2014). Creation of a computer self-efficacy measure: Analysis of internal consistency, psychometric properties, and validity. *Cyberpsychology, Behavior, and Social Networking*, *17*, 677–681.

Howard, M. C., & Jayne, B. S. (2015). An analysis of more than 1,400 articles, 900 scales, and 17 years of research: The state of scales in cyberpsychology, behavior, and social networking. *Cyberpsychology, Behavior, and Social Networking*, *18*, 181–187.

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, *31*, 313–323.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*, 187–200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*, 401–415.

Kline, P. (2014). *An easy guide to factor analysis*. New York, NY: Routledge.

Korobili, S., Togia, A., & Malliari, A. (2010). Computer anxiety and attitudes among undergraduate students in Greece. *Computers in Human Behavior*, *26*, 399–405.

Lam, L. T., & Li, Y. (2013). The validation of the R-victimisation scale (E-VS) and the E-bullying scale (E-BS) for adolescents. *Computers in Human Behavior*, *29*, 3–7.

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

Namlu, A. G., & Odabasi, H. F. (2007). Unethical computer using behavior scale: A study of reliability and validity on Turkish university students. *Computers & Education*, *48*, 205–215.

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*, 396–402.

Ozkan, S., & Koseler, R. (2009). Multidimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, *53*, 1285–1296.

Patil, V. H., Singh, S. N., Mishra, S., & Donavan D. T. (2007). Parallel analysis engine to aid determining number of factors to retain [Computer software]. Available from http://smishra.faculty.ku.edu/parallelengine.htm

Patil, V. H., Singh, S. N., Mishra, S., & Donavan D. T. (2008). Efficient theory development and factor retention criteria: A case for abandoning the 'eigenvalue greater than one' criterion. *Journal of Business Research*, *61*, 162–170.

Pawlikowski, M., Altstötter-Gleich, C., & Brand, M. (2013). Validation and psychometric properties of a short version of Young's Internet Addiction Test. *Computers in Human Behavior*, *29*, 1212–1223.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences*, *2*, 13–43.

Qin, H., Rau, P. L. P., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human–Computer Interaction*, *25*, 107–133.

Rainer, R. K., & Miller, M. D. (1996). An assessment of the psychometric properties of the computer attitude scale. *Computers in Human Behavior*, *12*, 93–105.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Tobias, S., & Carlson, J. E. (1969). Brief report: Bartlett's test of sphericity and chance findings in factor analysis. *Multivariate Behavioral Research*, *4*, 375–377.

van Braak, J., & Tearle, P. (2007). The computer attributes for learning scale (CALS) among university students: Scale development and relationship with actual computer use for learning. *Computers in Human Behavior*, *23*, 2966–2982.

Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*, 231–251.

Yang, T., Linder, J., & Bolchini, D. (2012). DEEP: design-oriented evaluation of perceived usability. *International Journal of Human–Computer Interaction*, *28*, 308–346.

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

Zoski, K. W., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement*, *56*, 443–451.

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

ABOUT THE AUTHOR

Matt C. Howard is a current PhD candidate at the Pennsylvania State University. His research interests include statistics and methodologies, workplace technologies, and employee training. Please visit MattCHoward.com for more information about the author, his publications, and helpful statistical guides.