JOURNAL OF MARKETING THEORY AND PRACTICE
https://doi.org/10.1080/10696679.2024.2385377

# A systematic review of exploratory factor analysis in marketing: providing recommended guidelines and evaluating current practices

Matt C. Howard[a] and Rachel O'Sullivan[b]

[a]Mitchell College of Business, University of South Alabama, Mobile, Alabama, USA; [b]University Advancement, Athens State University, Athens, Alabama, USA

**ABSTRACT**
We systematically review applications of exploratory factor analysis (EFA) in marketing since 2000. We show that most authors do not report EFAs in sufficient detail to determine whether appropriate practices were applied. When reported, most EFAs were conducted with practices that have been outdated for decades, resulting in inaccurate assessments of constructs. For instance, authors regularly utilize sample sizes too small to produce reliable estimates, apply inappropriate rotations, and use retention methods known to misidentify the number of factors. Our discussion highlights practices most in need of improvement, and we provide a checklist to ensure appropriate applications of EFA.

The foundation of empirical research is adequate measurement, whether through the creation of accurate sensors in the natural sciences or scales in the social sciences (Cohen et al., 2013; Hair et al., 2019; Stevens, 2012). If measures do not sufficiently represent intended constructs, empirical observations would not reflect phenomenon as they naturally occur, and subsequent inferences derived from these observations would be misleading. By utilizing poor measures, nonsignificant effects in the population could appear significant in a sample (i.e. Type I Error), and significant effects in the population could appear nonsignificant in a sample (i.e. Type II Error). Given the essential importance of proper measurement, an array of techniques has been developed to investigate the psychometric properties of measures. Exploratory factor analysis (EFA) is among the most popular of these techniques, and it is commonly applied in marketing due to the field's reliance on scales (Hair et al., 2019; Kamakura & Wedel, 2000; Peterson, 2000; Rossiter, 2002; Steenkamp & Maydeu-Olivares, 2023; Stewart, 1981).

EFA identifies a number of latent factors underlying a set of indicators, and it identifies the relation of each latent factor to each indicator – known as factor loadings (Goretzko et al., 2021; Howard, 2016, 2023; Watkins, 2018). For instance, a researcher may perform an EFA on a five-item service quality scale and a five-item customer satisfaction scale, discovering that four items solely relate to a first factor, four items solely relate to a second factor, one item relates to both factors, and one item relates to neither factor. Based on these results, it would be typical for the researcher to remove the latter two items to improve measurement adequacy, resulting in two four-item scales believed to represent independent constructs. Due to its data driven nature, EFA is believed to be apt at identifying whether a set of indicators adequately represent common and/or distinct constructs as intended by the applied scale(s) (Auerswald & Moshagen, 2019; Maskey et al., 2018; Nguyen & Waller, 2022). Likewise, because EFA assesses the relation of each factor with each indicator, it is also believed to be apt at identifying items that relate to multiple constructs (i.e. cross-loadings) (Brown, 2015; Li et al., 2020; Sass & Schmitt, 2010). Therefore, in addition to being a popular analysis, EFA is also an effective analysis.

More so than most other techniques, researchers must make several analytical decisions when conducting an EFA (Costello & Osborne, 2005; Fabrigar et al., 1999; Howard & Henderson, 2023). Choosing incorrect decisions can result in misleading EFA results, which could cause researchers to claim support for inaccurate measures. Using the prior example, an incorrectly performed EFA could cause the two problematic items to be included in the two produced scales, resulting in biased measurement tools. Ultimately, poorly

**CONTACT** Matt C. Howard MHoward@SouthAlabama.edu Mitchell College of Business, University of South Alabama, 5811 USA Drive S. Rm. 337, Mobile, AL 36688, USA

This article was originally published with errors, which have now been corrected in the online version. Please see Correction (http://doi.org/10.1080/10696679.2024.2395720)
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10696679.2024.2385377

© 2024 Taylor & Francis Group, LLC

performed EFAs could cause researchers to conduct subsequent analyses with inappropriate measures that misrepresent the studied phenomenon, resulting in incorrect assessment of hypotheses and misleading interpretations (Cohen et al., 2013; Hair et al., 2019). For these reasons, it is paramount that researchers understand and apply the correct decisions for performing an EFA.

Concerningly, three primary tensions are evident in the current marketing literature, stemming from changes in best practices for EFA since the last focused reviews of EFA in marketing (Peterson, 2000; Stewart, 1981). First, marketing researchers are likely applying outdated EFA approaches and perhaps engaging in practices now known to produce inaccurate findings, resulting in misleading insights into the substantive nature of their scales. For instance, modern guides now recommend the oblique rotations and a variety of factor retention methods (e.g. parallel analysis) should be applied when conducting EFA, but marketing researchers may still be utilizing orthogonal rotations and the Kaiser criterion. Second, several statistical and technological advances have been made since these reviews, including the creation of advanced technologies that can effortlessly conduct analyses that were previously restricted to only those with state-of-the-art computers. Aspects of EFA may not have been addressed in these original sources due to analytical or computing limitations at the time, leaving uncertainty regarding certain EFA decisions even after their publication. Because these technologies are now widespread, we expose readers to these advanced approaches for improved EFAs (e.g. machine learning), such that researchers can obtain more accurate EFA results. Third, researchers have increasingly recognized that proper EFA practices depend on the context, and certain EFA decisions may differ based on the field of study, such as factor loading interpretations (Howard, 2023; Howard & Henderson, 2023; Sakaluk & Short, 2017). Because this was not known at the time of these prior reviews, EFA best practices specific to marketing remain unrecognized.

To resolve these tensions in the literature, the current article reports a systematic review of EFA practices in marketing since 2000. We begin by discussing currently recommended EFA practices for psychometric investigations derived from relevant guides (e.g. Hair et al., 2019; Howard, 2016, 2023) and investigations (e.g. Goretzko & Bühner, 2020; Lim & Jahng, 2019). We briefly discuss the use of EFA for Harman's single-factor test and recent evidence that it is ineffective for assessing problematic common method bias (e.g. Fuller et al., 2016). Then, we report our systematic procedures and findings. We highlight certain widespread EFA practices in marketing that are particularly outdated or problematic, drawing researchers' attention to aspects of their methodological toolkits that could particularly benefit from updating. In our discussion, we use our systematic review to inform certain novel EFA recommendations that are specific to marketing, such as new

**Table 1.** Recommended EFA practices.

| Step | Recommendations |
|---|---|
| Step 0 – Review of Literature | Review the current literature for new developments to EFA. Review both statistical developments and software developments, as previously recommended but difficult to apply analyses may become easier with new programs. |
| Step 1 – Data Inspection | Ensure that sample sizes are appropriate based on estimates of common variance, such as the number of studied indicators and expected factors. We recommend the cutoffs of Rouquette and Falissard (2011). |
| | Perform typical assessments of data quality and approaches for data cleaning, such as recommended practices to ensure sufficient participant motivation and address missing data. Ensure that gathered samples are representative of the target population. |
| | Report Bartlett's test and KMO test. |
| Step 2 – Factor Extraction | Determine whether the investigation at hand calls for the analysis of total or common variance. Report the justification for studying either. |
| | If the study of total variance is desired, apply PCA. |
| | If the study of common variance is desired, apply PAF due to its less restrictive data assumptions. However, if model fit indices are desired, apply ML but recognize its data assumptions. |
| Step 3 – Factor Retention | Apply both visual scree plot analysis and parallel analysis. |
| | If these two approaches do not provide a firm factor retention decision, apply another supported method. We recommend the comparison data method due to its accuracy and ease of application. |
| Step 4 – Factor Rotation | Do not perform an orthogonal rotation. |
| | Perform an oblique rotation, such as oblimin or promax. |
| Step 5 – Factor Loading Interpretation | In an a priori manner, consider whether factor loadings are expected to be small, medium, or large based on the studied constructs. Then, use percentile values in Table 3 associated with these expectations as factor loading cutoffs. |
| | If there are no expectations for the size of factor loadings, apply our percentile cutoffs for moderate factor loadings. This would include a primary loading cutoff of .56, a secondary loading cutoff of .42, and a factor loading difference cutoff of .35. |
| | Describe the magnitude of resultant factor loadings based on the percentiles in Table 3. |
| Step 6 – Review of Reporting | Ensure that all recommended practices are clearly detained in the final EFA reporting. If needed, use supplemental materials or online repositories. |

guidelines for the interpretation of factor loadings. We conclude by providing a clear guide for best practices in conducting EFAs within marketing (Table 1).

By achieving these goals, the current article produces several implications for research and practice. First, we eliminate ambiguities in the current literature by providing guidance for all aspects of EFA, including those without widely accepted guidelines (e.g. factor loading interpretation). These efforts enable future analyses to be more accurate and easier to conduct, allowing researchers to better test their hypotheses and models. Second, while some reviewed recommendations are long-standing best practices, it cannot be guaranteed that researchers are utilizing these suggestions. We provide insights into whether the current marketing literature is overlooking well-known recommendations, which would suggest that greater efforts are needed by editors and reviewers to ensure proper statistical applications. Third, we identify specific studies that may have been particularly likely to produce misleading results by identifying problematic practices in prior applications of EFA in marketing research. Through using our systematic review database (Supplemental Material A), researchers could reinvestigate these past findings and perhaps discover the substantive nature of widely studied constructs may notably differ (e.g. number of factors) when reanalyzed using recommended techniques. Fourth, our review provides evidence that Harman's single-factor test is popular but inaccurate, thereby redirecting research efforts to analyses that are more effective in detecting common method bias. Together, we intend for the current article to be a clear guide for marketing researchers to perform improved EFAs, resulting in more accurate measurement and empirical research.

## Background

### *Exploratory and confirmatory factor analysis*

We first distinguish EFA from its most closely related analysis, confirmatory factor analysis (CFA), as the two analyses are indeed distinct and used for different purposes. When conducting an EFA, the researcher should have a general conceptualization of their studied constructs, including expectations regarding the number of dimensions and relations of each indicator to those dimensions; however, EFA does not specifically test a designated model, as it is instead a data-driven approach (Goretzko et al., 2021; Howard, 2016, 2023; Howard & Henderson, 2023; Stewart, 1981; Watkins, 2018). The analysis identifies the best model for the analyzed data, which includes the number of latent factors and the relation of each latent factor to the indicators. These relations include the association of indicators to their posited construct, known as factor loadings, and it includes the association of indicators to each other alternative construct, known as secondary loadings. This approach enables researchers to easily identify any unanticipated latent factors or secondary loadings, which is a clear benefit of the analysis.

Alternatively, the researcher is expected to have a firm conceptualization of their studied constructs and their relations when conducting a CFA (Brown, 2015; Harrington, 2009). The analysis specifically tests a designated model, and it assesses the extent that the modeled relations explain the underlying covariance between indicators in the studied dataset. Model fit indices are statistics utilized to determine whether the model adequately explains this covariance, and if adequately fitting, the researcher can then interpret the factor loadings and construct relations to understand the (inter)relations of indicators and latent factors. Because relations must be explicitly modeled to be tested via CFA, it is difficult to identify unanticipated latent factors or cross-loadings with the analysis. Nevertheless, the ability to assess the validity of a specific model is a sizable benefit of CFA, as researchers can determine whether their model and any competing models are adequate representations of the data (Brown, 2015).

EFA and CFA are similar in that they both assess the (inter)relations of indicators and latent factors; however, the two analyses differ in their approach, which causes them to produce differing considerations. The data-driven approach of EFA often enables researchers to identify problematic features of the studied indicators, such as cross-loadings, as these problematic features are rarely modeled when conducting CFA. For this reason, EFA is often applied when studying relatively untested measures, whether because they are newly created or recently adapted (Goretzko et al., 2021; Stewart, 1981; Watkins, 2018). It is also regularly used in the earlier phases of the scale development process. Some authors have called for EFA to be applied more broadly in empirical research, even when applying established measures, as researchers could better assess whether unexpected factors or loadings occurred in their studied sample (Howard, 2016, 2023; Howard & Henderson, 2023). Alternatively, CFA provides more direct inferences about a specific model, and researchers can assess whether their model of interest is an adequate explanation for the underlying structure of their data. This has caused authors to frequently apply CFA when studying established measures, as stronger expectations can be had for those measures (Brown, 2015; Harrington, 2009). CFA is also regularly applied in the latter phases of the scale development process. Together, EFA and CFA should be applied based on the needs of the researcher, and neither analysis should be considered superior to the other.

### *Psychometric investigations*

The primary focus of the current article is the use of EFA for psychometric investigation, which refers to applying the analysis to identify and substantively interpret the latent factors underlying a set of indicators (Hair et al., 2019; Peterson, 2000; Stewart, 1981). When performing an EFA for this purpose, five primary categories of decisions must be made, which are commonly categorized as (1) Data Inspection, (2) Factor Extraction, (3) Factor Rotation, (4) Factor Retention, and (5) Factor Loading Interpretation. We review each category below.

#### *Data Inspection*

Typical data inspection techniques should be applied when performing an EFA. For instance, researchers should utilize attention checks to ensure that participants were sufficiently motivated to accurately respond (Kung et al., 2018), and they should use modern approaches to address missing data (Newman, 2014). Researchers should also ensure that their sample is representative of their target population, which is essential for all empirical research. Many methodologists have commented on the frequent inadequacy of research samples, as convenient sources may not always generalize to intended target populations (Hanel et al., 2016; Stevens, 2011). For instance, it would be unlikely that results obtained from unemployed undergraduate students would generalize to employees currently in sales positions, as these students would have different experiences and perceptions than sales employees. Similarly, authors have discussed that samples are typically Western, educated, industrialized, rich, and democratic (WEIRD) (Cheon et al., 2020; Klein et al., 2022; Muthukrishna et al., 2020). If researchers intend for their samples to generalize beyond WEIRD contexts, they need to explicitly sample from those contexts. These considerations are true when using EFA, and the representativeness of the studied sample is essential when conducting the analysis.

Two data inspection techniques specific to EFA should also be applied when conducting the analysis, which are Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy. Bartlett's test determines whether the observed indicator correlation matrix is significantly different than an identity matrix (correlations of zeros), whereas the KMO test assesses the extent that each indicator is predicted by each other indicator (Lüdecke et al., 2021). Both analyses determine whether the data is suitable for EFA by testing whether sufficient interrelationships are evident between the indicators. When indicators share an insufficient amount of covariance, factor solutions may appear to be suitable, but they would instead explain little shared variance. In this case, the emergent factors would have little theoretical utility, and the researcher could unknowingly derive misleading theoretical inferences from the deceptive factors. For this reason, these two analyses are essential to perform in all applications of EFA (Costello & Osborne, 2005; Goretzko et al., 2021; Howard, 2023; Watkins, 2018), as researchers cannot otherwise detect when these instances may occur. A significant Bartlett's test and a KMO value above .60 indicate that the data is appropriate for EFA (Hair et al., 2019; Howard, 2016).

Additionally, researchers must obtain an appropriate sample size to conduct EFA, for which many recommendations are widespread in the current literature (Goretzko et al., 2021; Howard, 2023; Howard & Henderson, 2023; Maskey et al., 2018; Reio & Shuck, 2015; Sakaluk & Short, 2017; Watkins, 2018). The most popular recommendations are absolute guidelines and relative guidelines. Absolute guidelines recommend a fixed minimum number of participants, and the most popular absolute guidelines are 200 or 300 participants. Relative guidelines recommend a minimum number of participants based on the number of analyzed indicators in the EFA, of which 5 or 10 participants per indicator are the most popular recommendations.

Methodologists have shown, however, that the necessary sample size for EFA is not a fixed value, and it depends on more than the number of studied indicators (Costello & Osborne, 2005; Hogarty et al., 2005; MacCallum et al., 1999; Rouquette & Falissard, 2011). Instead, the minimum number of participants depends on the pattern of common variance shared by the studied indicators. While common variance can be observed by aspects such as communalities or factor loadings, these aspects are often difficult to predict before conducting the EFA. For this reason, authors have recommended that researchers should base their sample sizes on both the number of studied indicators and the expected number of emergent factors, as these two aspects together often correspond to the common variance of indicators (Howard, 2023; Howard & Henderson, 2023; Rouquette & Falissard, 2011). Few authors have provided sample size tables for EFA based on these aspects, but Rouquette and Falissard's

(2011) tables are among the most empirically supported because the authors conducted extensive simulations to arrive at their sample size recommendations. Based on their simulations, these authors recommend sample sizes that are larger than common absolute or relative guidelines, ranging from 250 (PCA, 45 indicators, and 2 factors) to 800 participants (EFA, 10 indicators, and 3 factors). Researchers should be aware that often larger sample sizes are necessary to obtain accurate EFA results in most applied contexts than those commonly recommended in the literature, and they should use Rouquette and Falissard (2011) tables to determine appropriate sample sizes for their EFAs.

### Factor Extraction Method

Three factor extraction methods are commonly discussed in extant guides: principal components analysis (PCA), principal axis factoring (PAF), and maximum likelihood (ML). The first of these, PCA, analyzes the total variance of indicators, whereas the latter two, PAF and ML, analyze the common variance of indicators. In earlier works of EFA (e.g. Harris, 1964; Jeffers, 1967; Lawley & Maxwell, 1962; Velicer, 1974), the implications of studying total or common variance were often debated, but more recent articles have less frequently discussed this distinction in depth. Instead, contemporary authors typically apply a factor extraction method without providing supporting justifications beyond an occasional citation, which is a problematic practice because the study of total or common variance poses differing theoretical implications.

The goal of PCA is to maximize the amount of explained variance in the indicators, which is why PCA involves the analysis of total variance rather than common variance (Hair et al., 2019). Because PCA analyzes total variance, it is often not considered a true factor analytic technique (Howard, 2016, 2023; Howard & Henderson, 2023). Instead, methodologists typically consider the study of total variance to be the analysis of components, whereas the study of common variance is the analysis of factors. Researchers should be careful with their language when discussing PCA for this reason. Further, the analysis of total variance does not differentiate between shared and unique variance, causing PCA results to represent a combination of both. Authors have suggested that PCA may be most applicable when specific and error variances are expected to represent a small portion of the studied indicators' total variance. PCA is also regularly applied when a researcher's objective is to summarize indicators into a minimum number of components for prediction purposes, as the analysis of total variance is particularly effective for data reduction (Fabrigar et al., 1999; Hair et al., 2019). Despite these differences between PCA and factor analysis, the two sets of analytic techniques produce similar results when the subsequent steps of the EFA process are performed correctly (Matsunaga, 2010; Trendafilov et al., 2013; Velicer & Jackson, 1990), emphasizing the importance of each EFA decision.

Alternatively, PAF and ML are considered factor analytic techniques because they assess common rather than total variance, which can be particularly useful when the researcher does not have knowledge or expectations about their indicators' specific and error variance (Hair et al., 2019). Regarding the distinction of PAF and ML, the former imposes fewer data assumptions, enabling it to provide accurate results in a wider range of contexts. The latter natively provides model fit indices (Costello & Osborne, 2005); however, Montoya and Edwards (2021) found that model fit is an ineffective approach to determine the number of factors to retain from EFA, drawing the benefits of model fit indices with ML into question. For this reason, we recommend the application of PAF for the study of common variance to avoid restrictive data assumptions (Beavers et al., 2019; Costello & Osborne, 2005; Watkins, 2018).

Together, PCA should be used when authors wish to study total variance, and PAF and ML should be used when authors wish to study common variance. If common variance is of interest, authors should use PAF when possible due to its less restrictive assumptions, but they should use ML if the provision of fit indices is necessary. Regardless of the approach, researchers should provide justification for their applied extraction method.

### Factor Retention Method

The researcher must determine the number of factors to retain from their EFA, and most methods to make this decision use eigenvalues (Courtney, 2013; Larsen & Warne, 2010; Patil et al., 2008). Each factor has an eigenvalue, which is the sum of squared factor loadings for that factor. As factor loadings represent the relation between a factor and an indicator, eigenvalues represent the amount of variance that a factor explains in the indicators. As the goal of EFA is to retain factors that explain a meaningful amount of variance in the indicators and exclude factors that do not, the goal of factor retention methods is to identify this distinction via eigenvalues.

The most popular factor retention method is the Kaiser criterion, but it is unfortunately among the least accurate factor retention methods (Braeken & Van Assen, 2017; Humphreys & Ilgen, 1969; Patil et al., 2008). The Kaiser criterion proposes that researchers should retain all factors with an eigenvalue above 1.00 and exclude all factors with an eigenvalue below 1.00. Simulation studies have repeatedly shown that the Kaiser criterion regularly overestimates the number of factors to retain, causing EFA solutions to include factors that do not explain meaningful amounts of variance in indicators (Braeken & Van Assen, 2017; Humphreys & Ilgen, 1969; Patil et al., 2008). Authors have also regularly commented that factors with eigenvalues of 1.01 and 0.99 are largely indistinguishable regarding the variance that they explain in indicators, and retaining one while excluding the other is arbitrary and counter-intuitive. Given these considerations, most modern guides for EFA recommend against utilizing the Kaiser criterion.

Another popular factor retention technique is the visual scree plot analysis (Macrosson, 1999; Roberson et al., 2014; Yong & Pearce, 2013). To perform this approach, eigenvalues are plotted sequentially on a chart from the first to last factor. The researcher then identifies the final "elbow" in the eigenvalues, or the final point that the eigenvalues significantly decrease in value. Factors before the elbow are retained, and those after the elbow are excluded. Figure 1 provides a demonstration of a visual scree plot analysis, showing an elbow in eigenvalues. Modern guides often recommend performing a visual scree plot analysis, but only in conjunction with other retention approaches due to its subjectivity (Maroof, 2012; Watkins, 2018). It can be difficult to determine exactly where the final elbow occurs when performing a visual scree plot analysis, causing some discrepancy in interpretations.

While less popular than the Kaiser criterion or visual scree plot analysis, parallel analysis is among the most supported factor retention techniques (Dinno, 2009; Ledesma & Valero-Mora, 2007; Lim & Jahng, 2019). When conducting a parallel analysis, the researcher performs an initial EFA, and a specified number of datasets (e.g. 10000) consisting of random numbers are generated with the same sample size and number of indicators as the conducted EFA. Separate EFAs are performed on each of these datasets of random numbers, and the eigenvalues of each factor from each dataset are recorded. These eigenvalues represent variance explained in spurious relations alone, and any meaningful eigenvalue in the EFA conducted on actual data should be significantly larger than these eigenvalues derived from random values. For this reason, researchers typically retain factors with eigenvalues above the 95th percentile of the eigenvalues derived from random values when performing a parallel analysis. In recent years, parallel analysis has become increasingly easy to apply, as websites now exist that generate the datasets and conduct the EFAs automatically (e.g. Patil et al.,

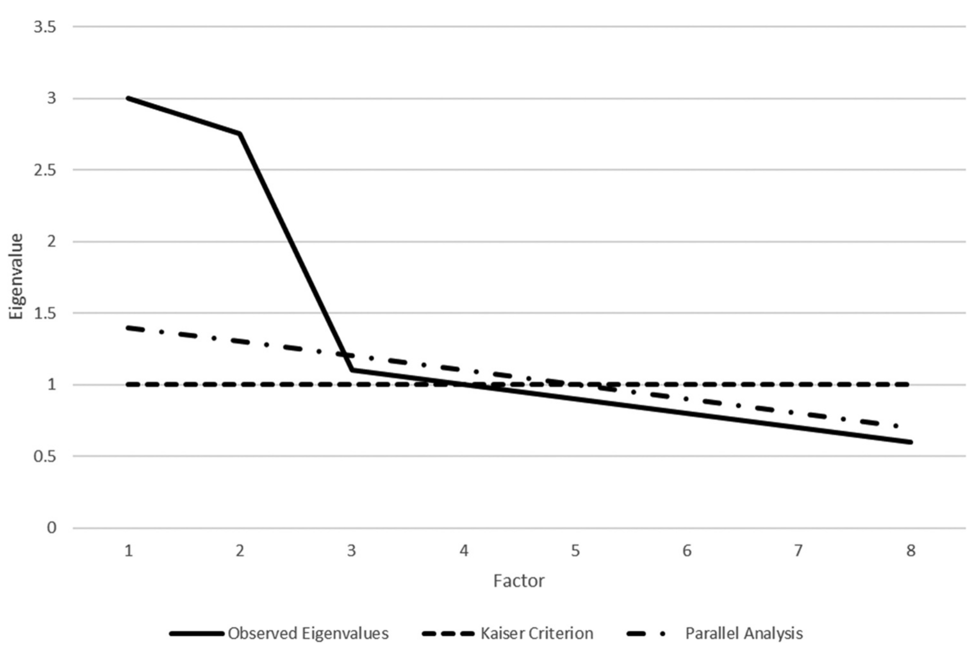


Figure 1. Visual representation of factor retention based on eigenvalues. Figure 1 illustrates the application of three different factor retention approaches. In this example, the Kaiser criterion recommends the retention of three factors, as the eigenvalues for the first three factors are above 1.00. The visual scree plot analysis recommends the retention of two factors, as the "elbow" occurs after the second factor; only the first two eigenvalues occur before the point of diminishing returns, and the subsequent eigenvalues explain much smaller amounts of variance in the indicators. The parallel analysis recommends the retention of two factors, as the eigenvalues of only the first two factors are larger than the parallel analysis eigenvalues.

2017). Due to its accuracy and ease of use, it is expected that parallel analysis will become more popular in the future.

A multitude of other factor retention approaches have also been supported to be more accurate than the Kaiser criterion and visual scree plot analysis, but many of these have yet to become widespread. Many of these other approaches currently require coding languages (e.g. R, Python) or sparsely used programs to apply. It is expected that these other approaches will become more popular in future research, as methodologists are continuously creating programs to ease the application of more accurate statistics. We recommend that researchers should monitor developments to the machine learning approaches (Goretzko & Bühner, 2020) for factor retention decisions, as significant support has been provided for its accuracy.

Together, we recommend that researchers should apply both visual scree plot analysis and parallel analysis, and they could provide further evidence for their retention decisions by also applying supported but less common approaches (e.g. machine learning). If the number of factors to retain is unclear with both visual scree plot analysis and parallel analysis, then researchers must supplement these two analyses with a third approach. As highlighted in our discussion, we recommend applying the comparison data method in this circumstance. The comparison data method provides accurate factor retention decisions (Auerswald & Moshagen, 2019), and a user-friendly program has been developed for its application (Courtney, 2013)

### *Factor Rotation Method*

When performing an EFA, the initial solution is a correct representation of the relations between the emergent latent factors and indicators, but an infinite number of potentially more interpretable alternative solutions are also correct representations. Factor rotations are defined statistical approaches to identifying a correct and more interpretable alternative solution, which can greatly aid the researcher in understanding the nature of their EFA results.

The decision of which family of rotation to apply often has larger implications for conducting an EFA than the exact rotation, and the two families of rotations are orthogonal and oblique (Maskey et al., 2018; Reio & Shuck, 2015; Sakaluk & Short, 2017). Orthogonal rotations do not allow emergent factors to be correlated, whereas oblique rotations allow (but do not force) emergent factors to be correlated. If a justification is provided for applying an orthogonal rotation, authors usually specify that they did not expect sizable correlations among their emergent factors, and therefore they did not choose a rotation that could model factor correlations. At the same time, if a justification is provided for applying an oblique rotation, authors usually specify that all emergent factors are correlated to some extent, and therefore they chose a rotation that could model these factor correlations and provide more accurate results. Relevant reviews and guides of EFA have recommended the application of oblique rotations, as modeling these correlations – even when minimal – is necessary for interpreting factor solutions (Brown, 2015; Goretzko et al., 2021; Howard, 2023; Howard & Henderson, 2023).

Of specific rotations, varimax is the most popular orthogonal rotation across many fields of study (Beavers et al., 2019; Costello & Osborne, 2005; Fabrigar et al., 1999; Ledesma et al., 2021; Watkins, 2018). The varimax rotation increases the variance in factor loadings, producing a greater dispersion between large and small factor loadings. Because the varimax rotation does not allow emergent factors to be correlated, it is infrequently recommended in modern guides, and it has been less often applied in recent years in the related fields of tourism and hospitality management (Howard & Henderson, 2023) and management (Howard, 2023).

Some authors have applied the varimax rotation when their studied constructs are highly correlated, and they believe that obtaining uncorrelated factor scores is necessary to address multicollinearity in subsequent analyses. Guides for addressing multicollinearity typically do not recommend this approach and/or explicitly recommend against it (Kraha et al., 2012; Lin, 2008; Sulaiman et al., 2021; Tu et al., 2004). Constructs are always correlated to some extent, even if relatively minor. While varimax produces orthogonal factors, the true nature of the represented constructs still includes this correlation, and the varimax rotation simply fails to model this shared variance. The results produced from the varimax rotation do not align with the true nature of the constructs, even if used in an attempt to address multicollinearity issues. For this reason, other approaches are recommended to address multicollinearity issues, such as theoretically driven model revision, instead of utilizing varimax rotation (Kraha et al., 2012; Lin, 2008).

The most popular oblique rotations are promax and oblimin in reviews of other academic domains (Costello & Osborne, 2005; Fabrigar et al., 1999; Ford et al., 1986; Henson & Roberts, 2006; Howard, 2023; Howard & Henderson, 2023). The promax rotation achieves a similar function to varimax (increase variance in factor loadings), but it allows emergent factors to be correlated. The oblimin rotation similarly differentiates large and small factor loadings, doing so by "minimize[ing] the correlations of loadings across factors" (Spicer, 2005, p. 188). Present research has yet to firmly demonstrate that promax or oblimin rotations produce more accurate results beyond the other, and the same is true for other oblique rotations. For this reason, we recommend the application of oblique rotations, but we do not specify a specific oblique rotation to apply (akin to other guides) (Beavers et al., 2019; Ledesma et al., 2021; Watkins, 2018).

### Factor Loading Interpretation

When performing an EFA, researchers must interpret primary factor loadings and cross-loadings. Primary loadings are the indicators' factor loadings on intended factors, whereas cross-loadings are the indicators' factor loadings on alternative factors. It is typical to retain indicators with large primary loadings and small cross-loadings while excluding indicators with small primary loadings and/or large cross-loadings (Ford et al., 1986; Peterson, 2000; Sass, 2010). For this reason, it is necessary to distinguish large from small when considering primary loadings and cross-loadings, and this designation should be determined before conducting the EFA.

Earlier guides discussed the interpretation of statistical significance when assessing primary loadings and cross loadings (Cattell, 1966; Kim & Mueller, 1978); however, more recent sources advise against this practice because statistical significance is influenced by more than the loading alone, and researchers observed that the meaningfulness of a factor loading is often independent from its statistical significance (Brown, 2015; Costello & Osborne, 2005; Ford et al., 1986). These sources instead recommend absolute cutoffs for primary loading and cross-loading interpretations. Unfortunately, a wide array of recommendations can be seen in recent reviews of EFA (Goretzko et al., 2021; Howard, 2016, 2023; Howard & Henderson, 2023; Maskey et al., 2018; Reio & Shuck, 2015; Sakaluk & Short, 2017; Watkins, 2018). Common recommendations for primary loading cutoffs range from .32 to .70, whereas common recommendations for cross-loading cutoffs range from .30 to .70. Perhaps equally problematic, authors rarely provide justifications for applying their chosen cutoff beyond a citation, leaving it unclear why any cutoff may be better or worse than another. Therefore, no primary loading or cross-loading cutoff is dominant or most supported in the present literature.

Recent years have also seen an increase in authors considering the difference between indicators' primary loadings and cross-loadings (Topcu-Uzer et al., 2021; Tuck & Thompson, 2023; Vicente et al., 2020). Absolute guidelines can result in problematic indicator retention decisions. For instance, if a researcher applied primary loading and cross-loading cutoffs of .40, they would retain an item with a primary loading of .41 and a cross-loading of .39, although a .02 discrepancy in loadings is likely not a meaningful difference. Among the few recommendations for the difference between primary loadings and cross-loadings, Howard (2016) suggested a cutoff of .20 based on a review of EFA in the human-computer interaction literature.

These differences in interpretations and wide ranges of cutoffs may be due to the relatively arbitrary nature of factor loadings. That is, factor loadings do not have an inherent characteristic that signifies large or meaningful values. For instance, while a factor loading of .50 is larger than a factor loading of .40, researchers cannot decisively label either as large or meaningful based on the nature of factor loadings themselves. For this reason, Howard (2023) recently adopted prior approaches for developing correlational effect size benchmarks (Bosco et al., 2015; Gignac & Szodorai, 2016; Orth et al., 2022) to craft new factor loading guidelines for management research. The author conducted a systematic review of EFA in management to identify benchmark ranges for primary loadings, cross-loading, and loading differences based on percentile estimates of previously reported findings, and they offered the ranges of very small (≤$10^{th}$ percentile), small ($10^{th}$-$33^{rd}$ percentile), moderate ($33^{rd}$-$66^{th}$ percentile), large ($66^{th}$-$90^{th}$ percentile), and very large (≥$90^{th}$ percentile). These ranges actualized the strength of factor loadings accepted in the field of management, as they were created from factor loadings of published EFAs. The author suggested that authors could require their indicators to produce factor loadings that meet the cutoff for moderate values, but these ranges also enable researchers to select more liberal or conservative cutoffs based on their applications. Therefore, these factor loading benchmarks provided clarity to the interpretation of factor loadings in management.

It cannot be guaranteed that standards of management generalize to marketing, as the former may be more conservative or liberal regarding the treatment of factor loadings. For this reason, we replicate this creation of factor loading benchmarks using our review of marketing. In doing so, we provide a table of factor loading percentiles to provide much-needed clarity for the interpretation of factor loadings, as no cutoffs are currently dominant in the literature. Future researchers can use these percentiles as cutoffs, which is illustrated in our discussion.

### *Harman's single-factor test*

In addition to using EFA for psychometric investigations, the analysis is also commonly applied to detect problematic amounts of common-method variance, known as Harman's single-factor test (Fuller et al., 2016; Howard, 2023). To conduct Harman's single-factor test, researchers perform an EFA on all studied indicators. If a one-factor solution emerges and/or more than 50% of the variance in indicators is explained by the first factor, the data is considered to contain a problematic amount of common-method variance. Recent investigations have shown that Harman's single-factor test is ineffective, as it too often provides supportive results for data with problematic common-method variance (Aguirre-Urreta & Hu, 2019; Fuller et al., 2016). Unfortunately, its appearance in prominent outlets of marketing even after the publication of these investigations suggests that the analysis is still broadly being applied despite its inefficacy. For this reason, we systematically investigate the frequency that researchers in marketing apply Harman's single-factor test and the frequency that their results were supportive. If the analysis excessively produced supportive results, then it could indicate that researchers claimed support for their data that may have contained problematic common-method variance. Such a discovery would both suggest that certain prior research is concerning and replication attempts are necessary to verify the validity of prior results. We also believe that this additional information is necessary to convince researchers to no longer apply Harman's single-factor test, as extant investigations have not been fully effective in curbing the application of the analysis.

### *Additional considerations*

The current article makes two additional considerations. First, it could be argued that any concerning result observed in our review may be irrelevant if practices are trending toward more appropriate applications of EFA. For this reason, we assess the relation of year with each aspect of EFA practices and results. If these relations are largely not statistically significant, then they would indicate that EFA practices have remained relatively consistent over the studied 20-year period, and any problematic practices are not trending toward appropriate applications.

Second, it could be argued that any concerning result may be irrelevant if researchers are replicating EFA results with CFA or measurement models (e.g. portion of structural equation modeling [SEM]). This is not a valid argument (see Howard, 2023; Howard & Henderson, 2023). CFAs and measurement models can produce misleading results if based on inappropriately conducted EFAs, and they can particularly produce the same biasing effects if conducted on the same sample as the EFA (Brown, 2015). Nevertheless, we assess the frequency that authors conduct subsequent CFAs and measurement models on the indicators retained from the EFA, and we assess the frequency that these analyses are conducted on the same or different samples. By doing so, we assess whether EFAs are largely being replicated.

## Method

We review the applications of EFA in premier journals of marketing, as deriving insights from premier outlets can provide heightened benefits for future research. Articles from these sources often produce an intensified influence on the current literature, and providing corrective recommendations for these articles can help ensure that researchers adopt our recommended practices. It is also believed that any problematic practices in these more-selective outlets would be even more widespread within other less-selective marketing journals, and therefore readers should be particularly sensitive to any concerns observed in these outlets.

To identify relevant outlets, we utilized the Academic Journals Guide by the Chartered Association of Business Schools (2018), which categorizes business journals into separate tiers by discipline. We chose to review all journals ranked 4* or 4 within the Marketing section, which included the outlets of *Journal of Consumer Psychology, Journal of Consumer Research, Journal of Marketing, Journal of Marketing Research, Journal of the Academy of Marketing Science, Marketing Science, International Journal of Research in Marketing, and Journal of Retailing*. We used Publish or Perish 8 to perform Google Scholar searches in November 2023. These searches were restricted from the year 2000 to the present, as our intent was to provide a review of modern EFA practices. In each of our eight selected outlets, we performed the six searches of "Exploratory Factor Analysis," "EFA," "Principal Components Analysis," "PCA," "Harman's Single-Factor Test," and "Harman's One-Factor Test" (quotations included). After removing duplicates, these 48 total searches produced an initial list of 615 articles to be coded.

In each phase of the coding process, two coders coded sources until sufficient interrater agreement was met for a set of 20 articles (Cohen's $\kappa$ or ICC $\geq .80$) (Gisev et al., 2013). Once met, the primary coder coded the remaining sources. In the first phase of the coding process, the two coders coded the number of EFAs reported in each source. This resulted in the reduction of our article list from 615 to

474 sources. Our search terms often appeared in the reference section of the excluded sources, which caused them to appear in our searches but not report an EFA. In a small number of instances, the authors reported an EFA as a demonstration in a methodology-focused article, which we did not code and instead excluded. EFAs reported for this purpose are typically utilized for their unique characteristics or subpar results, and analytical decisions in these methodology articles often differ from typical empirical investigations. Therefore, excluding these articles provided a clearer depiction of modern EFA practices in marketing.

In the second phase of the coding process, we coded each EFA for the characteristics below, resulting in the coding of exactly 1,000 EFAs from 474 sources. For each characteristic below, the aspect was coded as "Unknown" if it could not be determined from the article.

### Sample size

The sample size was recorded. If the sample size of the EFA was not explicitly stated, we recorded the sample size of the associated study.

### Number of indicators

The number of initially included indicators was recorded.

### Number of retained factors

The number of retained factors was recorded.

### Bartlett's test and KMO test

Whether the authors performed Bartlett's test or KMO test were each recorded as "Yes" or "No."

### Factor analytic method

The applied factor analytic method was recorded as "PCA," "PAF," "ML," or "Other."

### Factor retention method

The factor retention method was recorded as "Kaiser," "Scree," "Parallel," or "Other." Multiple factor retention methods could have been recorded for an EFA.

### Factor rotation method

The factor rotation method was recorded as "Orthogonal," "Oblique," or "None." The exact type of factor rotation method was also recorded as "Varimax," "Promax," "Oblimin," or "Other."

### Factor loading cutoffs

The primary factor loading cutoff, cross-loading cutoff, and difference between the primary factor loading and cross-loading cutoff was recorded.

### Factor loadings

The smallest primary loading, largest cross-loading, and smallest difference between a primary loading and cross-loading was recorded. Whether factor loadings were reported were also coded as "Yes" if all factor loadings were reported, "Partially" if some but not all were reported, or "No" if none were reported.

### Purpose

The purpose of the EFA was coded as "Psychometric Investigation" or "Harman's Single-Factor Test." Psychometric investigations assessed the factor structures underlying a set of indicators, whereas Harman's single-factor test assessed common-method variance.

### Harman's single-factor test cutoff

For Harman's single-factor tests, the applied cutoff was recorded as "50% of Variance in First Factor," "One Factor Solution," or "Both."

### Harman's single-factor test fail

Whether Harman's single-factor test was failed was recorded as "Yes" or "No."

### Year

The publication year of the article was recoded.

### Subsequent CFA

Whether a CFA or measurement model (either covariance-based SEM or partial least squares SEM; Hair et al., 2020) was conducted on the retained items from the EFA was coded as "Yes" or "No." If so, whether the authors utilized the same sample as the EFA was coded as "Same," and whether the authors used a different sample was coded as "Different."

## Results

Our final systematic literature review database is provided in Supplemental Material A. We first assessed whether any article was an outlier regarding the number of reported EFAs, such that we could ensure that no one article disproportionately affected our results. Following modern recommendations (Aguinis et al., 2013), we calculated z-scores for the number of EFAs each article reported. The most EFAs reported in the articles were 23 (z-score = 8.50), 22 (z-score = 8.09), 16 (z-score = 5.65), and 13 (z-score = 4.43). Six articles reported the next-highest number of EFAs (12; z-score = 4.02). Because the two articles that reported the most EFAs had significantly larger z-scores than the following articles, we removed these two articles from our analyses, resulting in a final database of 955 EFAs and 472 articles. Of these, 869 EFAs were psychometric investigations, whereas 86 were Harman's single-factor tests.

We also conducted a series of sensitivity analyses, which are analyses utilizing alternative approaches to ensure that our analytical decisions did not cause the current results. First, we calculated alternative results including the outlier articles discussed above, which are provided in Supplemental Material B. Second, we coded the EFAs exactly as they were reported, and these coding results were utilized in our primary analyses. For instance, some articles specified that they used PAF for their third EFA, but they may have not made this specification for their first or second EFAs. In this case, we coded the extraction method of the first and second EFAs as "Unknown," and the third EFA as "PAF." Some readers, however, may disagree with this approach, as they may believe that it could be safely assumed that the first and second EFAs likewise utilized PAF. For this reason, Supplemental Material C provides results of analyses wherein we assumed that all EFAs reported within an article utilized the same approach if not stated otherwise. In the example above, all three EFAs would be coded as "PAF." Third, our primary analyses report results for each EFA, causing some articles to appear in our dataset more than others. In Supplemental Material D, we report results with the EFAs coding averaged together for each article, such that each article appears in the dataset once. Fourth, we reconducted our analyses separated by outlet, which would be of interest to readers interested in the practices of specific outlets. These are provided in Supplemental Material E.

Across these sensitivity analyses, inferences obtained from the results remained consistent with our primary analyses, and no major deviations were seen for any aspect of modern EFA practices. These findings strongly support the robustness of our primary results, and they indicate that our findings were not solely due to our analytical decisions. Regarding the analyses of outlets, we did not observe sizable differences in reporting practices, and no outlet appeared to habitually report better or worse practices. The studied outlets seem similar in their standards and expectations for EFA, suggesting that researchers of marketing broadly should take note of our recommendations rather than only those appearing in a subset of outlets.

**Table 2.** Frequency of analytical decisions for EFA in marketing.

| | Frequency |
|---|---|
| Bartlett's Test | 29 (3%) |
| KMO Test | 44 (5%) |
| PAF | 53 (6%) |
| ML | 36 (4%) |
| PCA | 309 (36%) |
| Other | 2 (0%) |
| Unknown | 469 (54%) |
| Kaiser Alone | 134 (15%) |
| Scree Alone | 26 (3%) |
| PA Alone | 3 (0%) |
| Other Alone | 1 (0%) |
| K&S | 29 (3%) |
| K&PA | 2 (0%) |
| K&O | 2 (0%) |
| S&PA | 11 (1%) |
| S&O | 1 (0%) |
| PA&O | 0 (0%) |
| K&S&PA | 2 (0%) |
| K&S&O | 2 (0%) |
| K&PA&O | 1 (0%) |
| S&PA&O | 0 (0%) |
| K&S&PA&O | 0 (0%) |
| Unknown | 655 (75%) |
| Orthogonal | 171 (36%) |
| Varimax | 165 (35%) |
| Other Orthogonal | 1 (0%) |
| Oblique | 88 (19%) |
| Oblimin | 19 (4%) |
| Promax | 36 (8%) |
| None | 4 (1%) |
| Unknown | 207 (44%) |

All categories above are exclusive with the exception of factor rotations. The categories of orthogonal and oblique also include the frequencies of respective specific rotations within these categories of rotations. For instance, the provided frequency of orthogonal rotations includes the frequency of varimax rotations.

### *EFA for psychometric investigation*

Tables 2–4 provide our coding results. We first discuss data inspection procedures. Very few authors reported the results of either Bartlett's test (3%) or KMO test (5%), indicating that marketing researchers rarely report screening their data before conducting EFA. Regarding sample size, the median number of participants per EFA was 232, with a median participant-to-item ratio of 20. While the former was closer to small absolute guidelines for sample size, the latter exceeded even the largest widespread relative guidelines (Goretzko et al., 2021; Howard, 2023; Howard & Henderson, 2023; Maskey et al., 2018; Sakaluk &

**Table 3.** Median sample size of EFAs separated by number of indicators and factors.

| | Indicators | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Factors | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40+ |
| 1 | 236 (142) | 120 (39) | 272 (10) | - | - | - | - | - |
| 2 | 235 (64) | 246 (51) | 249 (12) | 341 (10) | - | - | - | - |
| 3 | 320 (14) | 227 (50) | 217 (17) | 224 (17) | - | - | - | - |
| 4 | - | 200 (13) | 212 (24) | 302 (15) | 117 (6) | - | - | 371 (10) |
| 5 | - | - | 249 (11) | 360 (7) | 140 (10) | - | - | 301 (5) |
| 6 | - | - | - | 292 (14) | 274 (6) | - | - | 214 (8) |
| 7 | - | - | - | - | - | - | 202 (9) | - |
| 8+ | - | - | - | - | 199 (8) | - | 447 (9) | 319 (15) |

Median sample sizes listed in each cell, and number of observations used to estimate the median sample size is provided in each cell within parentheses.

**Table 4.** Percentile estimates of EFA analytical decisions and results.

| | Sample Size | Number of Indicators | Sample Size to Indicator Ratio | Primary Loading Cutoff | Secondary Loading Cutoff | Difference Cutoff | Smallest Primary Loading | Largest Secondary Loading | Smallest Difference | Total Variance Explained |
|---|---|---|---|---|---|---|---|---|---|---|
| n | 850 | 805 | 798 | 98 | 50 | 0 | 308 | 153 | 138 | 320 |
| 10 | 100 | 4 | 4 | .40 | .25 | - | .43 | .21 | .04 | 57 |
| 20 | 126 | 6 | 7 | .40 | .25 | - | .50 | .29 | .11 | 62 |
| 30 | 162 | 7 | 11 | .40 | .25 | - | .54 | .32 | .14 | 66 |
| 33 | 172 | 8 | 12 | .40 | .30 | - | .56 | .34 | .16 | 68 |
| 40 | 195 | 9 | 15 | .44 | .30 | - | .60 | .36 | .20 | 70 |
| 50 | 232 | 11 | 20 | .50 | .30 | - | .63 | .39 | .27 | 72 |
| 60 | 292 | 15 | 28 | .50 | .30 | - | .66 | .41 | .32 | 75 |
| 66 | 320 | 18 | 35 | .59 | .30 | - | .69 | .42 | .35 | 76 |
| 70 | 352 | 20 | 39 | .60 | .30 | - | .70 | .44 | .36 | 78 |
| 80 | 464 | 25 | 59 | .60 | .36 | - | .73 | .47 | .42 | 83 |
| 90 | 781 | 36 | 119 | .70 | .50 | - | .80 | .50 | .47 | 88 |

The first row indicates the number of EFAs utilized to calculate the percentiles. The first column indicates the percentile values. For instance, the second row reports the 10th percentile for each column.

Short, 2017; Watkins, 2018). To assess whether marketing authors base their sample sizes on the number of indicators and/or factors, we conducted a regression analysis with sample size as the dependent variable and the number of indicators and factors as the independent variables.[1] Neither the number of indicators ($\beta = -.08$, $p = .14$) nor factors ($\beta = .07$, $p = .15$) had statistically significant relations with sample size. Table 3 presents median sample sizes separated by the number of indicators and factors. No median exceeded the sample size suggestions of Rouquette and Falissard (2011). Therefore, marketing researchers appear to not base their sample sizes for EFA on guidelines derived from estimates of common variance, as sample sizes did not correspond to proxies of common variance (number indicators and factors) and did not meet cutoffs based on common variance.

Determining a factor extraction method is the second step of EFA. It was most common for marketing researchers to not report their extraction method (54%). When reported, PCA (36%) was much more common than PAF (6%), ML (4%), or other extraction methods (~0%). This finding indicates that marketing researchers are much more interested in the analysis of total variance rather than common variance when performing EFA.

Determining a factor retention approach is the third step of EFA. It was again most common for marketing researchers to not report their factor retention approach (75%). Of reported approaches, it was most common for

[1]In conducting this analysis, we rescaled significant outliers for all three variables similar to the approach described above. Specifically, we calculated z-scores and rescaled large values to the next highest typical value based on the distribution of observations. This included rescaling six sample sizes to a value of 13,327 (*z-score* = 2.89), five numbers of indicators to a value of 93 (*z-score* = 3.96), and one number of factors to a value of 14 (*z-score* = 4.65).

researchers to apply the Kaiser criterion alone (15%), followed by the visual scree plot analysis alone (3%) and the Kaiser criterion and visual scree plot analysis together (3%). No other approach or combination of approaches appeared in more than one percent of EFAs. This finding indicates that marketing researchers are not following factor retention guidelines that have been widely recommended for decades (Dinno, 2009; Goretzko & Bühner, 2020; Ledesma & Valero-Mora, 2007; Lim & Jahng, 2019).

Determining a factor rotation method is the fourth step of EFA. We only assessed articles that did not report a one-factor solution when calculating estimates for this step. It was likewise most common for authors not to report their factor rotation approach (44%). Orthogonal rotations (36%) were more likely to be applied than oblique rotations (19%). Varimax was almost the sole orthogonal rotation applied (35%), as the remaining one percent solely included authors that did not specify their oblique rotation. More variety was seen in applied oblique rotations. Promax (8%) was the most common oblique rotation, which was followed by Oblimin (4%) and authors not specifying their applied oblique rotation (7%). Much like factor retention decisions, these findings suggest that marketing researchers are also not applying factor rotation guidelines that have been widely recommended for decades (Brown, 2015; Goretzko et al., 2021; Howard, 2023; Howard & Henderson, 2023; Maskey et al., 2018; Sakaluk & Short, 2017).

Interpreting factor loadings is the fifth step of EFA. Only 11% of articles reported an a priori primary factor loading cutoff; only 6% reported an a priori secondary factor loading cutoff; and no article reported a cutoff for the difference between primary and secondary loadings. Of articles that reported a primary factor loading cutoff, the most common were .40 (35%), .50 (23%), and .60 (17%). Retained indicators in reported EFAs almost always met the smallest of these cutoffs, as the 10$^{th}$ percentile of primary factor loadings was .43 and the median was .63. Of articles that reported a secondary loading cutoff, the most common were .25 (28%) and .30 (44%). Retained indicators in reported EFAs much less often met these cutoffs compared to primary loading cutoffs. The 90th percentile for largest secondary loading was .50, and the median largest secondary loading was .39. The 10th percentile for the smallest difference was .04, whereas the median smallest difference was .27. These results together indicate that an array of factor loading cutoffs are currently used in the marketing literature, and researchers typically retain items with strong primary loadings but also strong secondary loadings.

### *EFA for Harman's single-factor test*

We provide results regarding four aspects of Harman's single-factor tests: sample size, applied cutoffs, variance explained by first factor, and whether the test was passed. The median sample size for Harman's single-factor tests was 220, with a median participant-to-indicator ratio of 11. It was most common for researchers not to specify a cutoff (57%). A one-factor solution and/or the first factor explaining 50% or more variance was the most common cutoff (21%); a one-factor solution was the second most common cutoff (12%); and the first factor explaining 50% or more variance was the third most common cutoff (11%). Less than half of researchers directly stated the amount of variance explained by the first factor (48%). Of those that did, the first factor explained 27% of variance on average (*S.D.* = 7.77). No author found a one factor solution and/or the first factor to explain more than 50% of variance. Therefore, no researcher ever reported failing Harman's single-factor test in the studied marketing outlets.

### *Additional analyses*

We performed analyses with year as the sole predictor of each EFA discussed above. These results are provided in Table 5. The relations of year with the aspects were not statistically significant in 22 of 29 tests ($p > .05$). Of the significant relations, recent researchers were more likely to report their factor retention decisions ($\beta = -.03$, $p = .01$) and much more likely to apply the Kaiser criterion ($\beta = .04$, $p < .01$), whereas they were less likely to perform visual scree plot analyses ($\beta = -.04$, $p = .03$). Recent researchers were also more likely to apply the varimax ($\beta = .03$, $p = .04$) and promax rotations ($\beta = .05$, $p = .02$). Lastly, recent researchers reported using smaller primary ($\beta = -.31$, $p < .01$) and secondary loading cutoffs ($\beta = -.32$, $p = .02$).

We assessed the frequency that EFAs were replicated in subsequent CFAs or measurement models, and we assessed how often the CFAs or measurement models were conducted on the same or different samples. Sixty-two percent of EFAs included a subsequent CFA or measurement model. Of these, half were conducted with the same sample as the EFA, whereas half were conducted with a different sample as the EFA.

## Discussion

Adequate measurement is essential for accurate tests of hypotheses and advancement of theory. Due to the reliance of the survey methodology in marketing, EFA is a primary and particularly important analysis in the

**Table 5.** Results of regression analyses with year as sole predictor.

| | Coefficient | p | | Coefficient | p |
|---|---|---|---|---|---|
| **Psychometric Investigations** | | | | | |
| 1) Bartlett's Test | .04 | .19 | 2) KMO Test | .02 | .39 |
| 3) Sample Size | .02 | .60 | 4) Participant-to-Indicator Ratio | .06 | .09 |
| 5) Unknown Factor Extraction | −.01 | .21 | 6) PCA | .00 | .75 |
| 7) PAF | .02 | .32 | 8) ML | .02 | .37 |
| 9) Unknown Factor Retention | −.03 | .01* | 10) Kaiser Criterion | .04 | <.01** |
| 11) Visual Scree Plot Analysis | −.04 | .03* | 12) Parallel Analysis | .02 | .62 |
| 13) Other Factor Retention | −.01 | .83 | 14) Unknown Factor Rotation | −.02 | .12 |
| 15) Orthogonal Rotation | .02 | .07 | 16) Varimax | .03 | .04* |
| 17) Oblique Rotation | .00 | .77 | 18) Oblimin | .03 | .28 |
| 19) Promax | .05 | .02* | 20) Primary Loading Cutoff Provided | .02 | .22 |
| 21) Primary Loading Cutoff | −.31 | <.01** | 22) Smallest Primary Loading | .06 | .30 |
| 23) Secondary Loading Cutoff Provided | .01 | .57 | 24) Secondary Loading Cutoff | −.32 | .02* |
| 25) Largest Secondary Loading | −.14 | .09 | 26) Smallest Factor Loading Difference | .15 | .09 |
| **Harman's Single-Factor Test** | | | | | |
| 27) Sample Size | −.02 | .87 | 28) Participant-to-Indicator Ratio | −.03 | .82 |
| 29) First Factor Variance | .27 | .09 | | | |

Linear regression analyses were conducted for continuous outcomes, whereas binomial regression analyses were conducted for dichotomous outcomes.
* p < .05** p < .01

field, as it provides insights into the suitability of measures for representing studied constructs. Given its importance, the first goal of the current article was to review modern recommended practices for conducting EFA, such that comprehensive guidelines could be provided to modern researchers in marketing. The second goal of the current article was to review current EFA practices in the marketing literature since 2000 to determine areas of potential concern and avenues for improvement, such that marketing researchers could better understand which current practices need correcting. From our review of 1,000 EFAs from 474 sources, the current article identified several areas for improvement regarding current uses of EFA in the marketing literature, which can now be addressed to produce a more robust field of research with more accurate studies. We review our findings and provide suggestions regarding each step of the EFA process below. Table 1 summarizes our recommended practices.

## *Data inspection*

Very few researchers report conducting Bartlett's test or KMO test, which is a concerning practice in the marketing literature. When indicators share little common variance, EFA results may still appear meaningful, and issues may remain undetected by researchers; however, these results would provide little explanatory value regarding the interrelations underlying a set of indicators, and emergent latent factors would explain little in the indicators themselves (Hair et al., 2019; Lüdecke et al., 2021). In this case, theoretical insights that could be obtained from the emergent factors would be significantly hampered, stifling the meaningful progression of theory. It is possible – if not likely – that obtained factor structures in many prior studies were interpreted as substantive, but they were instead misleading depictions of the studied construct(s). Therefore, researchers should make greater efforts to conduct these two analyses, such that more confidence can be obtained in the theoretical relevance of the EFA results.

Regarding sample size, the median number of participants was on the smaller end of common absolute guidelines, but it was on the larger end of common participant-to-indicator guidelines (Goretzko et al., 2021; Maskey et al., 2018; Reio & Shuck, 2015; Sakaluk & Short, 2017; Watkins, 2018). Because researchers in marketing (on average) conduct their EFAs on a smaller number of indicators than other fields of study, such as tourism and hospitality management (Howard & Henderson, 2023) or management (Howard, 2023), they may believe that it is permissible to use sample sizes that are small by absolute guidelines. This is not the case. Analyzing a small number of indicators may call for larger sample sizes depending on underlying common variance, which is the largest known determinant of necessary sample sizes for EFA (Howard, 2023; Howard & Henderson, 2023; Rouquette & Falissard, 2011). Due to the importance of common variance for determining sample size, researchers have recommended basing EFA sample sizes on the ratio of indicators to factors, as this ratio is partially reflective of common variance. Our results also showed, however, that the number of indicators or latent factors did not significantly relate to sample size, suggesting that researchers are not basing sample sizes on either. Likewise, no median sample size for any combination for the number of indicators and latent factors met the cutoffs of Rouquette and Falissard (2011), which are based on simulations with varying common variance.

Thus, marketing researchers appear to not base their sample sizes on relevant characteristics of their data and likely underpower their EFAs.

Future researchers should recognize that sample sizes for EFA should be based on estimates of common variance, and the ratio of indicators to factors is likely the easiest estimate of common variance to estimate a priori. We recommend that marketing researchers should utilize the sample size recommendations of Rouquette and Falissard (2011), which provide sample size recommendations based on the number of indicators and expected factors. By utilizing these recommendations, researchers can ensure that their EFAs are sufficiently powered, and they can have greater confidence in the accuracy and robustness of their analyses.

### Factor extraction

Scholars of marketing strongly prefer the study of total variance rather than common variance, as PCA was applied almost three-and-a-half times more often than both PAF and ML together. Neither approach is more correct than the other because the analysis of total or common variance should be based on the theoretical rationale at hand. Nevertheless, we provide three recommendations for factor extraction in marketing based on our observations.

First, even more common than applying PCA, it was most common for researchers not to specify their factor extraction method. This practice is concerning. Greater attention has been paid in recent years to the importance of replication, which relies on the complete reporting of methods and analyses (Maxwell et al., 2015; Open Science Collaboration, 2015). Because differing results (with differing implications) are obtained by applying different factor extraction methods, many EFA results in marketing cannot be replicated. Future researchers should be more complete in their reporting. While page counts are limited, researchers could report their complete EFA procedures and results in supplemental materials or online repositories, such as the Open Science Framework (osf.io). By doing so, a more robust field of study can be obtained.

Second, researchers rarely provide a justification or rationale for their factor extraction approach. While neither the study of total nor common variance may be more correct than the other, it is still expected that researchers provide a justification for studying one beyond the other. Readers should refer to our conceptual distinctions between the study of total or common variance provided above, and readers should also reference prior sources that also discuss this distinction (e.g. Jain & Shandliya, 2013; Preacher & MacCallum, 2003). Likewise, even if studying common variance, researchers should provide justifications for the use of PAF, ML, or another approach. We recommend the use of PAF due to its fewer statistical assumptions.

Third, researchers have increasingly associated differing terminology with the study of total variance vs. common variance. Particularly, researchers consider the emergent dimensions from the analysis of total variance to be components or composites, whereas they consider the emergent dimensions from the analysis of common variance to be factors. Many researchers recognize this distinction when applying analogous analyses, such as covariance-based SEM vs. partial least squares SEM (Chin, 1998; Haenlein & Kaplan, 2004; Hair et al., 2019), but this terminology is not seen in discussions of EFA. We recommend that researchers should adopt this terminology from analogous discussions, which may cause greater attention to be paid to the conceptual differences between PCA, PAF, and ML.

### Factor retention

Most marketing researchers are not abiding by modern recommendations for factor retention. No factor retention approach was specified for three-fourths of EFAs. In these cases, it is unknown whether recommended procedures were applied, and it cannot be assured that the correct number of dimensions were obtained from the EFA. Equally concerning, researchers cannot replicate prior findings as described by the authors when the factor retention approach is not reported, leaving it unclear whether a large portion of published EFAs can be replicated.

When reported, the Kaiser criterion alone was the most popular factor retention approach. It was almost five times more often applied than the second most popular factor retention approach, and our supplemental analyses also supported that it is even more popular in recent years. For decades, researchers have recommended against using the Kaiser criterion alone, as it is the most inaccurate of the widely applied factor retention approaches (Braeken & Van Assen, 2017; Howard & Henderson, 2023; Humphreys & Ilgen, 1969; Patil et al., 2008). The Kaiser criterion routinely overestimates the number of factors that should be interpreted, and it is almost certain that many prior EFAs retained too many dimensions in their factor solutions.

Beyond the Kaiser criterion alone, only the visual scree plot analysis and Kaiser criterion with visual scree plot analysis were applied in more than one percent of EFAs. While visual scree plot analyses are a useful supplement to other approaches, they suffer from their subjective nature. Few – if any – recently published guidelines have recommended the sole use of the visual scree plot analysis or in conjunction with the Kaiser criterion alone (Goretzko & Bühner, 2020; Maroof, 2012; Watkins, 2018), and a large number of researchers may only be allocating nominal efforts to ensuring that the appropriate number of factors are retained from their EFAs.

These findings are quite concerning. Appropriate factor retention approaches ensure that the correct number of dimensions are obtained from the EFA. Utilizing subpar approaches can result in the identification of an incorrect number of emergent dimensions, ultimately causing misleading interpretations for the substantive nature of constructs and associated theory. Of all steps of the EFA process, factor retention may be the step that necessitates the largest changes in marketing research. Moving forward, we strongly urge researchers to utilize modern factor retention approaches. We particularly recommend the application of both visual scree plot analysis and parallel analysis. Simulation studies have supported that parallel analysis is among the most accurate factor retention approaches (Dinno, 2009; Ledesma & Valero-Mora, 2007; Lim & Jahng, 2019), and it can be conducted with freely available websites (e.g. Patil et al., 2017). When these two methods do not provide a conclusive factor retention decision, we recommend that researchers should also use the comparison data method, as prior studies have supported its accuracy (Auerswald & Moshagen, 2019; Courtney, 2013). This approach compares the observed eigenvalues to the eigenvalues of sequentially generated datasets with an increasing number of latent factors, and the number of factors that produce the most similar eigenvalues is retained. Courtney (2013) provides a resource that enables the comparison data method to be used in SPSS, enabling easier application of the method. Together, future researchers should see the application of visual scree plot analysis and parallel analysis as minimum guidelines for applying EFA, and they should apply other recommended approaches, such as the comparison data method, if these two methods do not provide a clear factor retention decision.

### Factor rotation

Like factor retention, marketing researchers are not abiding by modern recommendations regarding factor rotation approaches. No factor rotation was specified in almost half of EFAs. Again, this causes uncertainty regarding whether recommended procedures were applied, causing doubts regarding the validity and replicability of results. When reported, orthogonal rotations were applied almost twice as often as oblique rotations. Orthogonal rotations do not allow emergent latent factors to be correlated. Because all factors are expected to be correlated to some extent, resultant EFA solutions with orthogonal rotation do not accurately model the factors underlying a set of indicators (Maskey et al., 2018; Reio & Shuck, 2015; Sakaluk & Short, 2017). Therefore, the interpretation of orthogonally rotated factor solutions can be misleading, potentially causing researchers to provide inaccurate theoretical insights.

We recommend that marketing researchers should move away from orthogonal rotations, and they should instead apply oblique rotations. Oblique rotations allow emergent factors to be correlated, enabling resultant models to more accurately represent the variance of indicators and provide more accurate tests of theory (Brown, 2015; Costello & Osborne, 2005; Goretzko et al., 2021; Howard, 2023; Howard & Henderson, 2023). When marketing researchers apply oblique rotations, they appear to prefer promax in favor of oblique or other rotations, but no justification is currently widespread for applying any oblique rotation beyond another. Therefore, we do not provide a specific recommendation beyond the application of orthogonal rotations.

### Factor loading Interpretation

Few marketing researchers specify an a priori cutoff for primary or secondary loadings, and none consider a cutoff for the difference between primary and secondary loadings. When a primary loading cutoff was considered, researchers were generally split between cutoffs of .40, .50, and .60. When a secondary cutoff was considered, researchers were also split between cutoffs of .25 and .30. These findings suggest that marketing researchers may be presently uncertain regarding how to interpret their factor loadings, given the wide range of applied factor loading cutoffs. This uncertainty may have potentially caused undesirable indicators to be retained and desirable indicators to be excluded in final factor solutions from prior EFAs.

To address this tension, we propose a new approach to identifying factor loading cutoffs rather than arbitrarily selecting values. Factor loadings do not have intrinsic properties that define them as small, medium, or large (and thereby meaningful or not meaningful). While statistical significance tests can be conducted for factor loadings, authors have repeatedly shown that these are ineffective in interpreting the relations of indicators and latent factors (Brown, 2015; Costello & Osborne, 2005; Ford et al., 1986). Instead, we follow trends in identifying correlational effect size benchmarks to develop factor loading guidelines, which was also recommended in recent reviews of EFA in tourism and hospitality management (Howard & Henderson, 2023) and management (Howard, 2023). Several researchers have recognized that interpretations of correlations should be derived from the relative distribution of correlations in empirical research. Perhaps the historically most popular of these efforts is Cohen (1992) who provided a qualitative review of empirical research to derive correlational benchmarks. More recently, several authors have taken a more systematic approach to develop these benchmarks by conducting systematic reviews, cataloging correlational effect sizes, and providing estimates of percentile ranges for correlation magnitudes (Bosco et al., 2015; Gignac & Szodorai, 2016; Orth et al., 2022). Because these benchmarks are derived via systematic procedures, they are believed to be more theoretically sound and empirically supported. These recent efforts have also observed that correlation magnitudes differ across fields of study. As different fields study different phenomena, the typical sizes of their effects also differ.

To develop factor loading guidelines for marketing, we use Table 4. This table provides percentiles for the smallest primary loading, largest secondary loading, and smallest difference between the two. Following percentile ranges for correlation benchmarks (Bosco et al., 2015), we recommend that researchers should interpret factor loadings below the 10th percentile as very small, between the 10th and 33rd percentiles as small, between the 33rd and 66th percentiles as moderate, between the 66th and 90th percentiles as large, and above the 90th percentile as large. When applied to primary loadings, this would indicate that loadings below .43 are considered very small, loadings between .43 and .56 are small, loadings between .56 and .69 are moderate, loadings between .69 and .80 are large, and loadings above .80 are very large. When applied to secondary loadings, this would indicate that loadings below .21 are considered very small, loadings between .21 and .34 are small, loadings between .34 and .42 are moderate, loadings between .42 and .50 are large, and loadings above .50 are very large. When applied to primary and secondary loading differences, this would indicate that differences below .04 are considered very small, differences between .04 and .16 are small, differences between .16 and .35 are moderate, differences between .35 and .47 are large, and differences above .47 are very large.

To use these guidelines as a priori cutoffs, researchers could consider whether their studied constructs are expected to produce small, moderate, or large primary loadings, secondary loadings, and loading differences. For instance, if it is expected that the constructs would produce large primary factor loadings, the researcher could then apply an a priori primary loading cutoff of .69, as this is the minimum value associated with large loadings; if it is expected that the constructs would produce small secondary factor loadings, the researcher could then apply an a priori secondary loading cutoff of .34, as this is the maximum value associated with small loadings. We recommend that, if researchers do not have a strong expectation for the magnitude of their loadings, then they should apply cutoffs for moderate factor loadings. This would include a primary loading cutoff of .56, a secondary loading cutoff of .42, and a factor loading difference cutoff of .35. By developing these cutoffs, the current article provides clarity in interpreting factor loadings, and it also provides a rationale for interpreting factor loadings in marketing. Now, factor loading interpretations can be based on observed frequencies of magnitudes in marketing, rather than vague interpretations of small, medium, and large.

### *Harman's single-factor test*

Recent authors have expressed strong doubts regarding the validity of Harman's single-factor test (Aguirre-Urreta & Hu, 2019; Fuller et al., 2016). These authors argue that the analysis does not effectively assess common method variance, and it may too often enable researchers to claim support for data that is indeed overly influenced by common method variance (i.e. false negative) (Aguirre-Urreta & Hu, 2019; Fuller et al., 2016). Two recent reviews of EFA in tourism and hospitality management (Howard & Henderson, 2023) and management (Howard, 2023) even found that researchers in these fields have never found common method variance to be flagged as problematic when applying the analysis, regardless of the research design.

The present review found that Harman's single-factor test is still widely applied in marketing research, even after the publication of certain relevant articles on its inefficacy (Aguirre-Urreta & Hu, 2019; Fuller et al., 2016). Likewise, we also found evidence in marketing that corresponds with prior research in tourism and hospitality management (Howard & Henderson, 2023) and management (Howard, 2023), as researchers have likewise never found common method variance to be flagged as problematic when applying Harman's single-factor test in marketing. Prior researchers may have overstated support for their research designs, and potentially inappropriate inferences may persist in the field of marketing taken from studies with excessive common method bias. Due to this concerning result, we recommend that researchers should desist applying the analysis. They should instead apply other approaches to assessing common method variance that has been recommended in several reviews (Kock et al., 2021; MacKenzie & Podsakoff, 2012). Namely, authors have supported the effectiveness of techniques that require the measurement of specific constructs for identifying and addressing common method bias, which requires a priori considerations before the collection of data (Kock et al., 2021; MacKenzie & Podsakoff, 2012). We specifically recommend the application of these approaches, which include the marker-variable technique and careful usage of control variables, as they are better able to detect common method bias than Harman's single-factor test.

### Additional considerations

We discuss four additional considerations from our results. First, the current review observed many problematic EFA practices, suggesting that prior studies may have produced misleading insights regarding the nature of indicators and factors. For instance, small sample sizes and the Kaiser criterion both produce overestimates for the number of underlying latent factors, which may cause researchers to interpret constructs as containing more dimensions than they do in actuality (Braeken & Van Assen, 2017; Humphreys & Ilgen, 1969; Patil et al., 2008). We recommend that researchers should replicate prior EFAs with recommended approaches. It is possible that most prior findings produce the same results when using recommended approaches, and these replication efforts could support the robustness of modern theory in marketing; however, it is also possible that many of these prior findings do not produce the same results, and novel insights could be obtained regarding the substantive nature of marketing constructs. Readers could refer to Supplemental Material A to identify EFAs performed with particularly problematic practices, as these prior findings may benefit the most from reanalysis.

Second and relatedly, the application of improper EFA practices may have produced inappropriate inferences regarding theory in prior marketing studies. These inferences may include the substantive nature of constructs themselves, such as their representative dimensions; however, it may also include the relations of these constructs (Costello & Osborne, 2005; Goretzko et al., 2021; Watkins, 2018). Improperly performed EFAs can cause inappropriate items to be included in representative scales, which may systematically bias observed results. In the process of reanalyzing prior EFAs performed with substandard practices, researchers should also reassess whether the resultant inferences produce alternative interpretations of theory. Similar efforts have been made in alternative fields of study, such as the study of work engagement in management as discussed by Howard (2023). Early studies on the predominant work engagement scale produced a factor structure that was difficult to replicate, causing subsequent authors to strongly question the representative dimensions of work engagement. In turn, this caused the study of work engagement to slow, as these authors had to reinvestigate the most central question of their domain – what exactly is work engagement? Ultimately, authors reconducted the original analyses using recommended practices to create sounder interpretations of theory, enabling a significantly improved understanding of work engagement. Researchers in marketing should keep these efforts in mind when reviewing prior EFAs, as similar courses of action may be necessary to advance our understanding of predominant marketing constructs.

Third, the current article reviewed modern recommended practices for EFA, but new practices are constantly being developed. We did not discuss the most cutting-edge developments in EFA due to their present difficulty of application. For instance, many newly created practices have little documentation to guide applications and/or must be conducted with statistical software that requires coding beyond the typical researchers' expertise. Nevertheless, marketing researchers should monitor new developments to both EFA and statistical software. We particularly recommend the utilization of machine learning approaches with EFA, particularly factor retention. Indeed, Goretzko and Bühner (2020) developed machine learning procedures to estimate the number of factors to retain with EFA, and they supported that it is quite accurate with making factor retention decisions. If the analysis is developed in a user-friendly program, it could prove to be useful in combination with parallel analysis.

Fourth, our supplemental analyses provided evidence that researchers should not ignore problematic EFA practices by providing certain justifications. Namely, our results showed that EFA practices have remained largely unchanged over the studied 20-year period, and certain problematic practices are becoming even more common over time. For instance, recent researchers were more likely to apply both the Kaiser criterion and varimax rotation. Our results also showed that many researchers do not replicate their EFA results with subsequent CFAs or measurement models. Of those that did conduct a CFA or measurement model, half were conducted on the same sample, which is known to provide little psychometric evidence beyond the EFA conducted on that sample (Howard, 2023; Howard & Henderson, 2023). Given the few changes over time, it is even more important for researchers to use recommended EFA practices

**Table 6.** Summary of current and recommended practices for exploratory factor analysis.

| Step | Current Practices | Recommended Practices |
|---|---|---|
| 1) Data Inspection | Sample sizes are often modest for EFA, with an average sample size of 232. | Researchers should follow guidelines based on expected common variance, which include minimum sample sizes much larger than 232. |
| | Representative samples are often used, and general data quality checks are regularly applied. | Researchers should continue using representative samples and general data quality checks. |
| | Data quality checks specific to EFA are rarely used. | Researchers should apply Bartlett's test and KMO test. |
| 2) Factor Extraction | Researchers most often fail to report their applied factor extraction method. | Researchers should always report their chosen factor extraction method, as it is essential for replication purposes. |
| | When reported, principal components analysis is applied without justification. | Researchers should apply the analysis most relevant to their needs and justify their decision. Principal components analysis should be applied for the study of total variance; principal axis factoring should be applied for the study of common variance; and maximum likelihood should be applied if the provision of fit indices are desired. |
| 3) Factor Retention | Researchers most often fail to report their factor retention approach. | Researchers should always report their chosen factor retention approach, as it is essential for replication purposes. |
| | When reported, the Kaiser criterion alone is most often used. | Researchers should apply visual scree plot analysis and parallel analysis, and they should also consider applying other supported but less common approaches (e.g. machine learning). |
| 4) Factor Rotation | Researchers most often fail to report their applied factor rotation. | Researchers should always report their chosen factor rotation, as it is essential for replication purposes. |
| | When reported, varimax is the most common rotation, which is an orthogonal rotation. | Researchers should use oblique rotations, such as oblimin or promax. |
| 5) Factor Loading Interpretation | Researchers most often fail to specify their primary factor loading cutoffs. | Researchers should always report their chosen factor loading cutoffs, and they should be decided in an a priori manner. |
| | When reported, a wide array of primary and secondary loading cutoffs have been applied with little justification for the use of any. | Researchers decide a priori whether they expect their primary loadings, secondary loadings, and primary-secondary differences to be small, moderate, or large. They should then utilize our factor loading percentiles in Table 4 to set their factor loading cutoffs. If no expectation is had, they should apply cutoffs for moderate loadings: a primary loading cutoff of .56, a secondary loading cutoff of .42, and a factor loading difference cutoff of .35. They should report the magnitude of their loadings relative to our percentile values. |

to prevent inaccurate measures and misleading inferences from deterring marketing research.

### *Limitations*

As with any article, certain limitations should be noted. It cannot be guaranteed that EFA practices observed in our reviewed journals generalize to the marketing field as a whole, but we specifically chose to review the selected premier marketing journals for two primary reasons. These sources are known to have higher standards than less-prestigious outlets. It is likely that any concerning EFA practices in these outlets may be even more widespread in the broader marketing literature, and reviewing these premier outlets can reinforce that the observed concerns are likely to be seen more broadly. Also, articles from these chosen sources are known to have a heightened impact on the field. It is hoped that a review of these particularly important sources would likewise have a heightened impact. Nevertheless, future researchers should consider replicating the current investigation with a different selection of marketing outlets.

Also, we applied modern recommended practices in conducting a systematic methodological review. We applied relevant and comprehensive search procedures, and we achieved sufficient interrater agreement for our coding categories. It is always possible, though, that mistakes were made during the coding process. For this reason, future researchers should conduct similar reviews to ensure that our insights are accurate and valid.

## Conclusion

The current article reviewed modern recommended practices for EFA. We showed that marketing researchers often stray from recommended practices, and our results can be useful in highlighting practices that may benefit the most from changing. Our results also provided new guidelines for interpreting factor loadings and factor loading cutoffs, providing even further clarity to EFAs in marketing research. Table 6 provides a summary of current practices in the marketing literature, along with recommended practices for conducting EFA. It is believed that our results can produce more accurate analyses moving forward, ultimately resulting in a more accurate field of study.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## References


ABS. (2018). *Academic Journal Guide*. Chartered Association of Business Schools.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Aguirre-Urreta, M. I., & Hu, J. (2019). Detecting common method bias: Performance of the Harman's single-factor test. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *50*(2), 45–70. https://doi.org/10.1145/3330472.3330477

Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, *24*(4), 468. https://doi.org/10.1037/met0000200

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2019). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation*, *18*(1), 6.

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431. https://doi.org/10.1037/a0038047

Braeken, J., & Van Assen, M. A. (2017). An empirical Kaiser criterion. *Psychological Methods*, *22*(3), 450. https://doi.org/10.1037/met0000074

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In J. R. Nesselroade (Ed.), *Handbook of multivariate experimental psychology* (pp. 131–203). Springer US.

Cheon, B. K., Melani, I., & Hong, Y. Y. (2020). How USA-centric is psychology? An archival study of implicit assumptions of generalizability of findings to human nature based on origins of study samples. *Social Psychological & Personality Science*, *11*(7), 928–937. https://doi.org/10.1177/1948550620927269

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295–336.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

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), 7.

Courtney, M. G. (2013). Determining the number of factors to retain in EFA: Using the SPSS R-menu v2 0 to make more judicious estimations. *Practical Assessment, Research & Evaluation*, *18*(1), 8.

Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*, *44*(3), 362–388. https://doi.org/10.1080/00273170902938969

Fabrigar, L. R., Wegener, D., MacCallum, R., & Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272. https://doi.org/10.1037/1082-989X.4.3.272

Ford, J. K., MacCallum, R., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*(2), 291–314. https://doi.org/10.1111/j.1744-6570.1986.tb00583.x

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research*, *69*(8), 3192–3198. https://doi.org/10.1016/j.jbusres.2015.12.008

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality & Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social & Administrative Pharmacy*, *9*(3), 330–338. https://doi.org/10.1016/j.sapharm.2012.04.004

Goretzko, D., & Bühner, M. (2020). One model to rule them all? Using machine learning algorithms to determine the number of factors in exploratory factor analysis. *Psychological Methods*, *25*(6), 776. https://doi.org/10.1037/met0000262

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, *40*(7), 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics*, *3*(4), 283–297. https://doi.org/10.1207/s15328031us0304_4

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*. Cengage Learning.

Hair, J., Jr., Howard, M., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hanel, P. H., Vione, K. C., & Voracek, M. (2016). Do student samples provide an accurate estimate of the general public? *PLOS ONE*, *11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354

Harrington, D. (2009). *Confirmatory factor analysis*. Oxford university press.

Harris, C. W. (1964). Some recent developments in factor analysis. *Educational and Psychological Measurement*, *24*(2), 193–206. https://doi.org/10.1177/001316446402400202

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*(3), 393–416. https://doi.org/10.1177/0013164405282485

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, *65*(2), 202–226. https://doi.org/10.1177/0013164404267287

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction*, *32*(1), 51–62. https://doi.org/10.1080/10447318.2015.1087664

Howard, M. C. (2023). A systematic literature review of exploratory factor analyses in management. *Journal of Business Research*, *164*, 113969. https://doi.org/10.1016/j.jbusres.2023.113969

Howard, M. C., & Henderson, J. (2023). A review of exploratory factor analysis in tourism and hospitality research: Identifying current practices and avenues for improvement. *Journal of Business Research*, *154*, 113328. https://doi.org/10.1016/j.jbusres.2022.113328

Humphreys, L. G., & Ilgen, D. R. (1969). Note on a criterion for the number of common factors. *Educational and Psychological Measurement*, *29*(3), 571–578. https://doi.org/10.1177/001316446902900303

Jain, P. M., & Shandliya, V. K. (2013). A survey paper on comparative study between principal component analysis (PCA) and exploratory factor analysis (EFA). *International Journal of Computer Science and Applications*, *6*(2), 373–375.

Jeffers, J. N. (1967). Two case studies in the application of principal component analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *16*(3), 225–236. https://doi.org/10.2307/2985919

Kamakura, W. A., & Wedel, M. (2000). Factor analysis and missing data. *Journal of Marketing Research*, *37*(4), 490–498. https://doi.org/10.1509/jmkr.37.4.490.18795

Kim, J., & Mueller, C. (1978). *Factor analysis: Statistical methods and practical issues*. Sage.

Klein, V., Savaş, Ö., & Conley, T. D. (2022). How WEIRD and androcentric is sex research? Global inequities in study populations. *Journal of Sex Research*, *59*(7), 810–817. https://doi.org/10.1080/00224499.2021.1918050

Kock, F., Berbekova, A., & Assaf, A. G. (2021). Understanding and managing the threat of common method bias: Detection, prevention and control. *Tourism Management*, *86*, 104330. https://doi.org/10.1016/j.tourman.2021.104330

Kraha, A., Turner, H., Nimon, K., Zientek, L. R., & Henson, R. K. (2012). Tools to support interpreting multiple regression in the face of multicollinearity. *Frontiers in Psychology*, *3*, 44. https://doi.org/10.3389/fpsyg.2012.00044

Kung, F. Y., Kwok, N., & Brown, D. J. (2018). Are attention check questions a threat to scale validity? *Applied Psychology*, *67*(2), 264–283. https://doi.org/10.1111/apps.12108

Larsen, R., & Warne, R. T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods*, *42*(3), 871–876. https://doi.org/10.3758/BRM.42.3.871

Lawley, D. N., & Maxwell, A. E. (1962). Factor analysis as a statistical method. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *12*(3), 209–229. https://doi.org/10.2307/2986915

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation*, *12*(1), 2.

Ledesma, R., Ferrando, P., Trógolo, M., Poó, F., Tosi, J., & Castro, C. (2021). Exploratory factor analysis in transportation research: Current practices and recommendations. *Transportation Research Part F, Traffic Psychology and Behaviour*, *78*, 340–352. https://doi.org/10.1016/j.trf.2021.02.021

Li, Y., Wen, Z., Hau, K. T., Yuan, K. H., & Peng, Y. (2020). Effects of cross-loadings on determining the number of factors to retain. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(6), 841–863. https://doi.org/10.1080/10705511.2020.1745075

Lim, S., & Jahng, S. (2019). Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods*, *24*(4), 452. https://doi.org/10.1037/met0000230

Lin, F. J. (2008). Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. *Quality & Quantity*, *42*(3), 417–426. https://doi.org/10.1007/s11135-006-9055-1

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84. https://doi.org/10.1037/1082-989X.4.1.84

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common method bias in marketing: Causes, mechanisms, and procedural remedies. *Journal of Retailing*, *88*(4), 542–555. https://doi.org/10.1016/j.jretai.2012.08.001

Macrosson, W. D. K. (1999). Scree plots, data structure, and random variance. *Psychological Reports*, *84*(2), 533–540. https://doi.org/10.2466/pr0.1999.84.2.533

Maroof, D. A. (2012). Exploratory factor analysis. In D. A. Maroof (Ed.), *Statistical methods in neuropsychology: Common procedures made comprehensible* (pp. 23–34). Springer Science + Business Media, LLC. https://doi.org/10.1007/978-1-4614-3417-7_4

Maskey, R., Fei, J., & Nguyen, H. O. (2018). Use of exploratory factor analysis in maritime research. *The Asian Journal of Shipping and Logistics*, *34*(2), 91–111. https://doi.org/10.1016/j.ajsl.2018.06.006

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, *3*(1), 97–110. https://doi.org/10.21500/20112084.854

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *The American Psychologist*, *70*(6), 487. https://doi.org/10.1037/a0039400

Montoya, A. K., & Edwards, M. C. (2021). The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. *Educational and Psychological Measurement*, *81*(3), 413–440. https://doi.org/10.1177/0013164420942899

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, educated, industrial, rich, and democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, *31*(6), 678–701. https://doi.org/10.1177/0956797620916782

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*(4), 372–411. https://doi.org/10.1177/1094428114548590

Nguyen, H. V., & Waller, N. G. (2022). Local minima and factor rotations in exploratory factor analysis. *Psychological Methods*, *57*(1), 167–167. https://doi.org/10.1080/00273171.2021.2011697

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Orth, U., Meier, L. L., Bühler, J. L., Dapp, L. C., Krauss, S., Messerli, D., & Robins, R. W. (2022). Effect size guidelines for cross-lagged effects. *Psychological Methods*, *29*(2), 421–433. https://doi.org/10.1037/met0000499

Patil, V. H., Singh, S. N., Mishra, S., & Donavan, D. T. (2008). Efficient theory development and factor retention criteria: Abandon the 'eigenvalue greater than one'criterion. *Journal of Business Research*, *61*(2), 162–170. https://doi.org/10.1016/j.jbusres.2007.05.008

Patil, V. H., Singh, S. N., Mishra, S., & Donavan, D. T. (2017). *Parallel analysis engine to aid in determining number of factors to retain using R* [computer software]. https://analytics.gonzaga.edu/parallelengine/

Peterson, R. A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters*, *11*(3), 261–275. https://doi.org/10.1023/A:1008191211004

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*(1), 13–43. https://doi.org/10.1207/S15328031US0201_02

Reio, T. G., & Shuck, B. (2015). Exploratory factor analysis: Implications for theory, research, and practice. *Advances in Developing Human Resources*, *17*(1), 12–25. https://doi.org/10.1177/1523422314559804

Roberson, R. B., III, Elliott, T. R., Chang, J. E., & Hill, J. N. (2014). Exploratory factor analysis in rehabilitation psychology: A content analysis. *Rehabilitation Psychology*, *59*(4), 429. https://doi.org/10.1037/a0037899

Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing*, *19*(4), 305–335. https://doi.org/10.1016/S0167-8116(02)00097-6

Rouquette, A., & Falissard, B. (2011). Sample size requirements for the internal validation of psychiatric scales. *International Journal of Methods in Psychiatric Research*, *20*(4), 235–249. https://doi.org/10.1002/mpr.352

Sakaluk, J. K., & Short, S. D. (2017). A methodological review of exploratory factor analysis in sexuality research: Used practices, best practices, and data analysis resources. *Journal of Sex Research*, *54*(1), 1–9. https://doi.org/10.1080/00224499.2015.1137538

Sass, D. A. (2010). Factor loading estimation error and stability using exploratory factor analysis. *Educational and Psychological Measurement*, *70*(4), 557–577. https://doi.org/10.1177/0013164409355695

Sass, D. A., & Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, *45*(1), 73–103. https://doi.org/10.1080/00273170903504810

Spicer, J. (2005). *Making sense of multivariate data analysis: An intuitive approach*. Sage.

Steenkamp, J. B. E., & Maydeu-Olivares, A. (2023). Unrestricted factor analysis: A powerful alternative to confirmatory factor analysis. *Journal of the Academy of Marketing Science*, *51*(1), 86–113. https://doi.org/10.1007/s11747-022-00888-1

Stevens, C. K. (2011). Questions to consider when selecting student samples. *The Journal of Supply Chain Management*, *47*(3), 19–21. https://doi.org/10.1111/j.1745-493X.2011.03233.x

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Stewart, D. W. (1981). The application and misapplication of factor analysis in marketing research. *Journal of Marketing Research*, *18*(1), 51–62. https://doi.org/10.1177/002224378101800105

Sulaiman, M. S., Abood, M., Sinnakaudan, S., Shukor, M., You, G., & Chung, X. (2021). Assessing and solving multicollinearity in sediment transport prediction models using principal component analysis. *ISH Journal of Hydraulic Engineering*, *27*(sup1), 343–353. https://doi.org/10.1080/09715010.2019.1653799

Topcu-Uzer, C., Randall, A. K., Vedes, A. M., Reid, D., & Bodenmann, G. (2021). We-ness questionnaire: Development and validation. *Journal of Couple & Relationship Therapy*, *20*(3), 256–278. https://doi.org/10.1080/15332691.2020.1805082

Trendafilov, N. T., Unkel, S., & Krzanowski, W. (2013). Exploratory factor and principal component analyses: Some new aspects. *Statistics and Computing*, *23*(2), 209–220. https://doi.org/10.1007/s11222-011-9303-7

Tu, Y. K., Clerehugh, V., & Gilthorpe, M. S. (2004). Collinearity in linear regression is a serious problem in oral health research. *European Journal of Oral Sciences*, *112*(5), 389–397. https://doi.org/10.1111/j.1600-0722.2004.00160.x

Tuck, A. B., & Thompson, R. J. (2023). The social media use scale: Development and validation. *Assessment*, 31(3), 617–36.

Velicer, W. F. (1974). A comparison of the stability of factor analysis, principal component analysis, and rescaled image analysis. *Educational and Psychological Measurement*, *34*(3), 563–572. https://doi.org/10.1177/001316447403400309

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some further observations. *Multivariate Behavioral Research*, *25*(1), 97–114. https://doi.org/10.1207/s15327906mbr2501_12

Vicente, E., Verdugo, M. A., Guillén, V. M., Martínez-Molina, A., Gómez, L. E., & Ibáñez, A. (2020). Advances in the assessment of self-determination: Internal structure of a scale for people with intellectual disabilities aged 11 to 40. *Journal of Intellectual Disability Research*, *64*(9), 700–712. https://doi.org/10.1111/jir.12762

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *The Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079