International Journal of Nursing Studies 51 (2014) 511–521

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

Review

On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use

Cadeyrn J. Gaskin [a,b], Brenda Happell [c,*]

[a] Gaskin Research, Melbourne, Victoria, Australia
[b] Central Queensland University, Institute for Health and Social Science Research, Centre for Mental Health Nursing Innovation, School of Nursing and Midwifery, Melbourne, Victoria, Australia
[c] Central Queensland University, Institute for Health and Social Science Research, Centre for Mental Health Nursing Innovation, School of Nursing and Midwifery, Rockhampton, Queensland, Australia

ARTICLE INFO

Article history:
Received 27 June 2013
Received in revised form 28 September 2013
Accepted 4 October 2013

Keywords:
Research design
Nursing research
Review literature as topic
Statistics
Statistics as topic
Factor analysis
Principal components analysis
Measurement

ABSTRACT

Exploratory factor analysis (hereafter, factor analysis) is a complex statistical method that is integral to many fields of research. Using factor analysis requires researchers to make several decisions, each of which affects the solutions generated. In this paper, we focus on five major decisions that are made in conducting factor analysis: (i) establishing how large the sample needs to be, (ii) choosing between factor analysis and principal components analysis, (iii) determining the number of factors to retain, (iv) selecting a method of data extraction, and (v) deciding upon the methods of factor rotation. The purpose of this paper is threefold: (i) to review the literature with respect to these five decisions, (ii) to assess current practices in nursing research, and (iii) to offer recommendations for future use. The literature reviews illustrate that factor analysis remains a dynamic field of study, with recent research having practical implications for those who use this statistical method. The assessment was conducted on 54 factor analysis (and principal components analysis) solutions presented in the results sections of 28 papers published in the 2012 volumes of the 10 highest ranked nursing journals, based on their 5-year impact factors. The main findings from the assessment were that researchers commonly used (a) participants-to-items ratios for determining sample sizes (used for 43% of solutions), (b) principal components analysis (61%) rather than factor analysis (39%), (c) the eigenvalues greater than one rule and scree tests to decide upon the numbers of factors/components to retain (61% and 46%, respectively), (d) principal components analysis and unweighted least squares as methods of data extraction (61% and 19%, respectively), and (e) the Varimax method of rotation (44%). In general, well-established, but out-dated, heuristics and practices informed decision making with respect to the performance of factor analysis in nursing studies. Based on the findings from factor analysis research, it seems likely that the use of such methods may have had a material, adverse effect on the solutions generated. We offer recommendations for future practice with respect to each of the five decisions discussed in this paper.

* Corresponding author at: Central Queensland University, Institute for Health and Social Science Research, Centre for Mental Health Nursing Innovation, School of Nursing and Midwifery, Rockhampton, Queensland, Australia.Tel.: +61 07 49232164; fax: +61 07 49306402.
E-mail address: b.happell@cqu.edu.au (B. Happell).

0020-7489/$ – see front matter.
http://dx.doi.org/10.1016/j.ijnurstu.2013.10.005

What is already known about the topic?

- Exploratory factor analysis is sophisticated statistical procedure, for which knowledge on how it should be applied continues to develop.
- Many researchers do not use evidence-based methods when generating factor analytic solutions.

What this paper adds

- The assessment shows that researchers publishing in high-impact nursing journals frequently used out-dated heuristics to inform their decision making when performing factor analysis.
- Practical recommendations are made for conducting factor analysis in circumstances commonly encountered in nursing research.

1. Introduction

Exploratory factor analysis (hereafter, factor analysis), with its origins in the early twentieth century (Spearman, 1904), has become an integral statistical method in the social, health, biological, and, sometimes, physical sciences (Cudeck, 2007). Significant progress in the development of factor analysis was achieved when advances in computer technology enabled the timely performance of the substantial volumes of calculations required in the rotation of factors (Kaiser, 1958; Neuhaus and Wrigley, 1954). Another consequence of computerisation is that factor analysis has became accessible to many people with minimal training in the use of this method and who may accept the output of specialised software at face value (Browne, 2001). Without adequate training, researchers may seek to duplicate the factor analytic procedures that they see published in other studies, rather than making informed decisions about what methods are most appropriate for their circumstances.

Conducting factor analysis requires researchers to make several decisions (e.g., how large the sample size needs to be, and what methods to use for extracting and rotating factors), many of which have meaningful influences on the solutions generated (Fabrigar et al., 1999). Over time, various *rules of thumb* have been proposed to guide the decision making of researchers. As factor analysis has developed, however, evidence-based recommendations have replaced many of the original heuristics. Reviews of the use of factor analysis, however, have shown that researchers (perhaps due to ignorance) continue to use traditional heuristics ahead of evidence-based recommendations (e.g., Conway and Huffcutt, 2003; Fabrigar et al., 1999; Watson and Thompson, 2006).

This paper focuses on five major decisions that are made in conducting factor analysis: (i) establishing how large the sample needs to be, (ii) choosing between factor analysis and principal components analysis (PCA), (iii) determining the number of factors to retain, (iv) selecting a method of data extraction, and (v) deciding upon the method of factor rotation. The purpose of this paper is threefold: (i) to review the literature with respect to each of these five decisions, (ii) to assess current practices in nursing research, and (iii) to offer recommendations for future practice. We assume that readers have a basic understanding of factor analysis, which has been covered elsewhere (see, for example, Hair et al., 2010; Preacher and MacCallum, 2003; Tabachnick and Fidell, 2012). We begin with a description of how we performed the assessment of current factor analysis procedures, before turning to the five decisions.

2. Assessment procedure

The assessment was conducted using papers published during 2012 in the 10 nursing journals with the highest 5-year impact factors according to the *Journal Citation Reports*® (Science and Social Sciences 2011 editions): Birth, Oncology Nursing Forum, International Journal of Nursing Studies, Cancer Nursing, Journal of Advanced Nursing, Worldviews on Evidence-Based Nursing, Research in Nursing & Health, American Journal of Critical Care, Nursing Research, and Midwifery. These journals were selected because the studies published within them have, arguably, the greatest influence in the nursing research community. The inclusion of these 10 journals was to gain a broad impression of how factor analysis is used in nursing research.

Papers were included in this assessment if they reported findings using factor analysis or PCA. Although factor analysis and PCA are separate data extraction procedures with their own models, they were both included in this study because PCA is frequently (and mistakenly) considered to be a form of factor analysis. Papers were excluded from the study if the factor analyses were reported as part of the method sections rather than in the findings. In such circumstances, researchers often provide limited information about how they performed factor analysis.

Of the 773 research papers published in the 10 journals in 2012, 54 factor analyses (and PCA) were reported in the findings sections of 28 papers. Of the 54 factor analysis (and PCA) solutions reported, one factor analysis solution was reported in each of 16 papers, two solutions were described in each of 8 papers, and three or more solutions (specifically, 3, 4, 7 and 8, solutions) were detailed in each of 4 papers.

From each of the solutions, the following data were collected: (a) how the adequacy of the sample size was established, (b) whether factor analysis or PCA was used and how researchers made the choice between the two models, (c) how researchers determined the number of factors to retain, (d) what method of data extraction was used (including the type of data and whether Pearson or polychoric correlation matrices were used), and (f) what methods of factor rotation were used.

3. Establishing the required sample size

3.1. Review of literature

Determining the sample size required for factor analysis is challenging. In the absence of empirical evidence, various heuristics were developed. These rules have focused on, for example, (a) the absolute minimum sample size required (e.g., 100 participants; Kline, 1986); (b) the ratio of participants to variables (e.g., at least 10:1; Nunnally, 1978); and (c) the ratio of variables to expected factors (e.g., at least 3:1 to 6:1; Cattell, 1978). Unfortunately, the findings of studies conducted since these heuristics were proposed do not support their use (e.g., Hogarty et al., 2005; MacCallum et al., 1999; Mundfrom et al., 2005; Velicer and Fava, 1998). Rather, the findings from these studies suggest that the accuracy of factor solutions is dependent on the magnitudes of communalities (estimates of shared variance) and factor loadings, the sizes of samples, and the degree of overdetermination (extent to which each factor is represented by a distinct set of items). McCallum et al., for example, showed that more accurate solutions are achieved with larger sample sizes ($\hat{\omega} = .15$), high communalities ($\hat{\omega} = .41$), and greater overdetermination ($\hat{\omega} = .11$). Similarly, Hogarty et al. emphasised the importance of high communalities and overdetermination in producing quality factor solutions. Sample sizes below 50, for example, can be adequate when factor solutions exhibit ideal simple structures with (a) high factor loadings (e.g., .80 and above) or (b) mid-range factor loadings (e.g., .40, .60) with few factors (three or less) and large numbers of items (e.g., at least 12 with factor loadings of .60 and two factors extracted; de Winter et al., 2009). Such conditions do not commonly occur in practice, however, which means larger sample sizes are typically required. Although none of these studies have produced easily-applied methods for determining sample size, researchers can leverage the information in these studies to build justifications for choosing particular sample sizes in their work.

3.2. Assessment of practices evident in nursing journals

Although estimating sample sizes required for factor analysis remains a challenging task, the findings of the assessment suggest that there has been limited appreciation of the inadequacy of simple ratios. Researchers reported using participants-to-item ratios ($n = 23$ solutions; 43%), participants-to-parameters ratios ($n = 8$; 15%), and minimum numbers of participants ($n = 11$; 20%) as heuristics for determining sample sizes. For 10 solutions, two or more of these methods were used in combination. For over half the solutions ($n = 31$, 57%), no methods of determining the adequacy of sample sizes were reported. These percentages exceed 100%, because of the use of multiple methods for some solutions.

Unfortunately, the effect of the use of these heuristics cannot be determined from a simple assessment of the information presented in the studies. Inferences may be drawn, however, from simulation research in this area. Evidence suggests that the number of items per factor and the magnitudes of both communalities and factor loadings can affect the required sample sizes to such an extent as to render estimations using ratios hopelessly inaccurate (MacCallum et al., 1999, 2001). The implication of this research for studies included in this assessment is that it casts doubt over the accuracy of the factor analysis solutions that have been produced. Poor estimation of sample size requirements for factor analysis can lead to the recruitment of too many or too few participants, which not only affects the quality of the solutions generated, but can also have adverse ethical and resource implications (Gaskin and Happell, 2013).

Table 1
Minimum sample size recommendations for factor analysis.

Variables per factor	Communalities		
	High	Wide	Low
4	500	900	1400
6	250	200	260
8	100	130	130

Adapted from Mundfrom et al. (2005).
Note. Minimum sample sizes provided for sample solutions with excellent congruence to population solutions. High: communalities ranging between .60 and .80; wide: communalities ranging between .20 and .80; low: communalities ranging between .20 and .40.

3.3. Recommendations for future practice

Researchers are strongly encouraged to draw upon the findings of simulations in which minimum sample size requirement have been determined using multiple criteria (e.g., MacCallum et al., 1999; Mundfrom et al., 2005). In Table 1, some minimum sample size requirements are provided, which take into consideration the magnitudes of the communalities and the extent of overdetermination.

4. Choosing between factor analysis and principal components analysis

4.1. Review of literature

Factor analysis and PCA are separate data extraction procedures, each with their own purpose and model (Schmitt, 2011; Thomson, 1939; Widaman, 2007). Some authors have argued that the purpose of factor analysis is to explain the correlations between items in terms of one or more latent factors, whereas the objective of PCA is to represent as much of the variance in as few components as possible (Thomson, 1939; Widaman, 2007). That is, the aim of factor analysis is to describe variables in terms of a smaller number of underlying dimensions, whereas the goal of PCA is data reduction (Bandalos and Boehm-Kaufman, 2009). With factor analysis, only the shared variance among items is analysed, which is achieved through inserting communalities into the diagonal of the correlation (or covariance) matrix. In contrast, with PCA, all the variance in the items is included in the analysis, with unity values (i.e., 1.0) retained in the diagonal of the correlation (or covariance) matrix.

The choice between factor analysis and PCA has been the subject of substantial debate in the literature (see, for example, Bandalos and Boehm-Kaufman, 2009; Widaman, 2007 for reviews). Some of the arguments advanced in favour of factor analysis include that (a) the presence of an error (or uniqueness) term in the factor analytic model means that this model more accurately represents data in the real world, which almost always includes random error (Bentler and Kano, 1990), (b) this model provides clearly testable hypotheses about the data on which it is based (Bentler and Kano, 1990; McArdle, 1990), and (c) this model provides parameter estimates that are generalisable beyond the batteries for which they were originally generated (Widaman, 2007). In contrast, some arguments supportive of the use of PCA include that (a) this method has greater computational efficiency, meaning that it can handle larger numbers of variables and provide faster solutions (Velicer and Jackson, 1990b), (b) methods of determining the number of factors to retain are more developed for PCA than for factor analysis (Velicer and Jackson, 1990a), (c) factor analysis can have factor indeterminacy issues (Schönemann and Steiger, 1978), and (d) this model produces similar solutions to factor analysis (Fava and Velicer, 1992; Velicer, 1976b, 1977).

With the continued development of factor analysis and computer processing power, many of the advantages of PCA have diminished over time. First, although PCA will always be computationally more efficient than factor analysis (McArdle, 1990), the power of modern computers means that most researchers should not need to take their hardware into consideration when choosing between PCA and factor analysis. Second, advances in research mean that improved methods for determining the number of factors to retain are available to researchers using factor analysis (see the section *Determining the Number of Factors to Retain* in this paper). Third, the problem of the indeterminacy of factor analysis scores is disputed (see, for example, Fabrigar and Wegener, 2012; McArdle, 1990). Fourth, despite strong similarities between the solutions generated using the two models (Fava and Velicer, 1992; Velicer, 1976b, 1977), notable differences have been found when there was possible over-extraction of factors (Velicer, 1977) and when one, or a combination, of the following conditions have occurred: low component loadings (.40), small sample sizes ($N = 80$), and few items per component (4:1; Fava and Velicer, 1992). Consistent with these findings, several researchers have shown PCA to be less accurate than factor analysis, especially when loadings are low (.40) and when there are few items per factor/component (i.e., 3:1; Snook and Gorsuch, 1989; Widaman, 1993). Non-zero component loadings in PCA are systematically inflated, whereas the bias in factor loadings can be positive, neutral, or negative (Snook and Gorsuch, 1989; Widaman, 1990, 1993). In simulations representative of conditions that typically occur in practice (i.e., variable non-zero loadings on oblique factors/components, differing numbers of items per factor/component), factor analysis has been shown to be more accurate than PCA (Widaman, 1993).

4.2. Assessment of practices evident in nursing journals

The findings of the assessment show a preference for PCA over factor analysis. PCA was used to extract over half of the solutions ($n = 33$, 61%) and factor analysis was used for the remaining solutions ($n = 21$, 39%). No reasons were given as to whether, or why, one method was preferred over the other.

For two main reasons, the preference for PCA and the absence of explanations for this choice is of concern. First, it suggests that researchers have not fully appreciated that the goals of, and models underlying, PCA and factor analysis differ. Second, under conditions commonly encountered in nursing studies, empirical evidence suggests that factor analysis is more accurate than PCA (Widaman, 1993). The implication of the (perhaps) overuse of PCA in studies published in nursing journals is that the accuracy of many of the solutions is likely to have been diminished.

4.3. Recommendations for future practice

Researchers are encouraged to appreciate the differences between these models and choose the one that is most suitable for their work. For most nursing projects, on both conceptual and empirical grounds, factor analysis will be the more appropriate choice.

5. Determining the number of factors to retain

5.1. Review of literature

Several methods of determining the number of factors to retain have been presented in the literature, including Bartlett's (1951, 1950) test, Kaiser's (1960) eigenvalues greater than one rule, Cattell's (1966) scree test, Velicer's (1976a) minimum average partial (MAP) rule, Horn's (1965) parallel analysis, the Hull method (Lorenzo-Seva et al., 2011), and Ruscio and Roche's (2012) comparison data (CD). Parallel analysis has consistently been shown to have higher levels of accuracy than Bartlett's test, the eigenvalues greater than one rule, and the scree test (Hubbard and Allen, 1987; Velicer et al., 2000; Zwick and Velicer, 1986). Since the development of parallel analysis, several improvements to this method have been proposed (e.g., Crawford et al., 2010; Glorfeld, 1995). Most recently, researchers have proposed that conducting parallel analysis using principal axis factoring with the 95th percentile rule improves the accuracy of this method (Green et al., 2012). The MAP rule has performed strongly in some studies (Zwick and Velicer, 1982, 1986), but not others (e.g., Ruscio and Roche, 2012), and is only appropriate for use with PCA, not factor analysis (Velicer, 1976a). Both the Hull and CD methods outperformed parallel analyses in the initial studies in which these methods were compared (Lorenzo-Seva et al., 2011; Ruscio and Roche, 2012). More research needs to be conducted on these new methods, however, especially comparing each of these methods and the latest revision of parallel analysis (Green et al., 2012).

Determining the number of factors to retain with ordinal data presents researchers with additional challenges. Several researchers have proposed ways of determining the number of factors to extract from ordinal data (e.g., Cho et al., 2009; Liu and Rijmen, 2008). Recently, Garrido et al. (2012) provided evidence that researchers should perform parallel analysis with polychoric (rather than Pearson) correlations, using PCA as the method of extraction and the mean eigenvalue criterion. Parallel analysis with Pearson correlations and the mean eigenvalue criterion also performed well when data were only moderately skewed (0.00 to ±1.00). When the data exhibited higher levels of skewness, using Pearson correlations generally produced inaccurate results, hence the preference for polychoric correlations. These authors also make the point that the use of PCA for determining the number factors does not mean that this method needs to be used in the actual extraction of factors. That is, researchers can use one data extraction method for determining the number of factors to retain and another method for performing the factor analysis.

5.2. Assessment of practices evident in nursing journals

Researchers generally used older methods of determining the number of factors/components to retain (e.g., the eigenvalues greater than one rule, scree tests), which have been shown to be less accurate (i.e., they result in the extraction of too many or too few factors) than newer methods (e.g., parallel analysis). In their studies, researchers used the eigenvalues greater than one rule ($n = 33$, 61%), scree tests ($n = 25$, 46%), the findings of prior studies ($n = 11$, 20%), interpretations of the factor solutions ($n = 9$, 17%), parallel analysis ($n = 5$, 9%), total variance explained ($n = 3$, 6%), and the MAP rule ($n = 1$, 2%). For 7 solutions (13%), no method was mentioned. For just under half the solutions ($n = 23$, 43%), one method was used to determine the number of factors/components to retain, and for an equivalent number of solutions ($n = 24$, 44%), combinations of two to four methods were used.

The widespread use of the eigenvalues greater than one rule, often in conjunction with scree tests, is inconsistent with empirical evidence that has been generated over the last three decades (Hubbard and Allen, 1987; Ruscio and Roche, 2012; Velicer et al., 2000; Zwick and Velicer, 1986). The implication of this finding is that researchers may be extracting too many or too few factors/components in their analyses. The eigenvalues greater that one rule consistently, and often substantially, overestimates the number of factors (Ruscio and Roche, 2012; Zwick and Velicer, 1982, 1986). The scree test generally performs better than eigenvalues greater that one rule, but is less accurate than parallel analysis (Hubbard and Allen, 1987; Velicer et al., 2000; Zwick and Velicer, 1986) and, with inexperienced researchers, the inter-rater reliability of interpretations of this test can be quite low (Crawford and Koopman, 1979). Unfortunately, more accurate methods of determining the numbers of factors to retain (e.g., parallel analysis) were rarely used in the nursing studies reviewed.

5.3. Recommendations for future practice

The recommendations for future practice are dependent on the type of data that are being analysed (e.g., nominal, ordinal, interval, or ratio; Stevens, 1946). Similar to researchers in the social sciences (Croasmun and Ostrom, 2011), those publishing in nursing journals commonly use Likert-type response formats as a means to collect data (Jakobsson, 2004). An example of a Likert-type response format is one with five options labelled *strongly disagree*, *disagree*, *neutral*, *agree*, and *strongly agree*. Our choice of the term *response format* here is deliberate; the more frequently-used term, *scale* (as in *Likert scale*), is ambiguous, because it can refer to the response format (e.g., a 5-point scale anchored with *strongly agree* and *strongly disagree*), a type of measurement (e.g., nominal, ordinal, interval, and ratio scales) and a collection of items (e.g., 20-item scales; Carifio and Perla, 2007). Likert-type response formats typically generate ordinal data, but not always (Allen and Seaman, 2007). If the response options are equidistant then interval or ratio data can be generated. Evidence suggests, however, that producing response formats with options that are equidistant is a challenging, and not necessarily intuitive, task (Woltz et al., 2012). Unless researchers are able to demonstrate that the options provided to participants are equidistant, making the assumption that the data is ordinal, rather than interval or ratio, is advisable.

When researchers are working with continuous (interval, ratio) data, the evidence supports the use of parallel analysis with principal axis factoring and the 95th percentile rule. When working with ordinal data, researchers should seriously consider using parallel analysis with PCA as the method of extraction and the mean eigenvalue criterion. If the data are only moderately skewed (0.00 to ±1.00) then either Pearson or polychoric correlations can be used for this analysis. With highly skewed data, however, the use of Pearson correlations typically produces inaccurate results, so polychoric correlations should be used in these situations.

Researchers should keep in mind that the data extraction method used to determine the number of factors to retain (i.e., principal axis factoring or PCA, if these recommendations are being followed), does not need to be used subsequently to generate factor analysis solutions. The recommendations for data extraction methods (see Section 6) should be used for this task.

Parallel analysis and polychoric correlations may be unfamiliar terms to some users of commercial statistical software, because options to use these methods have not been available until recently. For users of SPSS (IBM Corp, 2012), an extension package is available to download for free, which enables researchers to perform parallel analyses with polychoric correlations (Basto and Pereira, 2012; IBM Corp, 2013). Although navigating the various websites to download and install several software packages to expand the functionality of SPSS can be a little complicated, Courtney (2013) has simplified this task through the inclusion of a helpful step-by-step guide in his recent paper. For those who use SAS, from the Second Maintenance Release of SAS 9.3 the option to perform polychoric correlations has been available (SAS Institute, 2013). O'Connor (2000) has published syntax for performing parallel analysis with SAS (an updated version of this syntax is available through the internet address mentioned at the end of O'Connor's paper). The syntax can be modified to accommodate polychoric correlations (Cho et al., 2009). Specialist non-commercial software for factor analysis is also available (Lee, 2010; Lorenzo-Seva and Ferrando, 2006). In short, these recommendations can be followed using either commercial or non-commercial software.

6. Selecting a data extraction method

6.1. Review of literature

Several extraction methods have been developed to enable the estimation of factor analysis models, including ordinary least squares factor analysis (equivalent to minimum residuals, Harman, 1960, also known as unweighted least squares factor analysis, Lee et al., 2012), image factor analysis (Guttman, 1953; Jöreskog, 1962, 1969), maximum likelihood methods (e.g., Lawley, 1940; Rao, 1955), generalised least squares (Browne, 1973; Jöreskog and Goldberger, 1972), principal axis factoring (Thomson, 1934), and alpha factor analysis (Kaiser and Caffrey, 1965). These methods differ in their objectives (e.g., principal axis factoring maximises the variance accounted for), definitions of uniqueness (the variance in an item not shared with other items), and procedures for computing communality estimates and factor scores (Gorsuch, 1983; MacCallum et al., 2007). Due to the dissimilarities between these extraction methods, they yield different solutions, even when sample sizes are very large (MacCallum and Tucker, 1991). Therefore, researchers need to apply extraction methods that are most suitable for their circumstances.

Selecting the most appropriate extraction method requires researchers to familiarise themselves with the backgrounds of each method, the assumptions that underlie them, and the empirical evidence regarding the conditions under which they produce satisfactory solutions. For example, the maximum likelihood method and the generalised least squares method require multivariate normal observed variables (Jöreskog, 2007; Jöreskog and Goldberger, 1972; Lawley, 1940), whereas the other methods do not require this assumption. In practice, however, the assumption of multivariate normality can rarely be met (Jöreskog, 2007). Although this reality may limit the possible use of maximum likelihood and generalised least squares methods, the use of maximum likelihood methods may be justifiable in many situations where the normality assumption is violated (Fuller and Hemmerle, 1966). Furthermore, evidence from work with confirmatory factor analysis has shown that maximum likelihood methods can be fairly robust under conditions in which the assumption of normality has been violated and performs better than the generalised least squares method (Boomsma and Hoogland, 2001). Caution must be used in applying this finding, however. Although exploratory factor analysis can be considered a special case of confirmatory factor analysis, exploratory factor analysis has unique properties, which mean that evidence about one model does not necessarily translate to the other (Yuan et al., 2002).

Several studies have been undertaken to compare extraction methods. Initial comparative research indicated a slight superiority of maximum likelihood methods over principal axis factoring (Browne, 1968). More recent research, however, suggests that the superiority of one of these two methods over the other may depend upon the prevailing conditions (de Winter and Dodou, 2012). In the de Winter and Dodou study, principal axis factoring yielded marginally more accurate estimates of the population pattern for simple solutions (i.e., equal and salient loadings [0.6] on all factors, orthorgonal factors, and each item loading on only one factor) than maximum likelihood factor analysis. Principal axis factoring also performed marginally better than the maximum likelihood method in the recovery of weak factors (i.e., when the number of items on a factor were reduced by half or two-thirds). The maximum likelihood method, however, outperformed principal axis factoring when there were severe distortions to the solution, such as unequal loadings on factors (non-zero loadings alternated between .90 and .30), and especially under conditions where there were unequal loadings in conjunction with correlations between factors, and under-extraction of factors. Like principal axis factoring, the ordinary least squares method may be superior to maximum likelihood in recovering weak factors (having loadings of .45 or below; Briggs and MacCallum, 2003; MacCallum and Tucker, 1991). The ordinary least squares method performs better than the maximum likelihood method when there is: (a) model error (simulated as 150 minor factors, each contributing .2 or .3 of the total variance) and a moderate fit of the model to the data (root mean square error of approximation = .03–.049), (b) sampling error from small sample sizes ($N = 100$), or (c) both model and sampling error (Briggs and MacCallum, 2003).

One issue with applying this knowledge on data extraction methods to most nursing studies is that it is based on the underlying assumption that the variables hold continuous data (Wirth and Edwards, 2007). Performing factor analysis with ordinal data using data extraction methods best suited for continuous data can produce spurious findings (Bernstein and Teng, 1989). Unfortunately, theory and methods for factor analysis with ordinal data are not as well-developed as for continuous data (Jöreskog and Moustaki, 2001).

The two main approaches used in addressing this issue are (a) assuming there is an underlying response variable (e.g., Jöreskog, 1994; Katsikatsou et al., 2012) and (b) item response theory (e.g., Bartholomew et al., 2011; Bock and Moustaki, 2007). The underlying response variable approach is based on the assumption that normally distributed, continuous variables underlie the ordinal response options, whereas item response theory treats the ordinal response options as they are. Taking the underlying response variable approach, for example, the use of ordinary least squares estimations with polychoric correlations has shown promise as a method of factor analysing ordinal data (Forero et al., 2009; Lee et al., 2012). With ordinal data, polychoric correlations between items can be stronger than the Pearson correlations (as are used with continuous data), especially when items are skewed (Holgado-Tello et al., 2010). Polychoric correlation matrices can also produce a better fit to theoretical models than Pearson correlation matrices. There were a range of conditions, however, under which ordinary least squares (and diagonally weighted least squares) methods have performed poorly, such as when there are few items per factor, binary response options, low factor loadings, high skewness, and small sample sizes (Forero et al., 2009).

Recently, a pairwise maximum likelihood method has been proposed, which may have advantages over the least squares methods (Katsikatsou et al., 2012).

6.2. Assessment of practices evident in nursing journals

Because there was a clear preference for PCA over factor analysis in the studies included in this assessment, it followed that PCA was the most commonly used method of data extraction. The methods used for data extraction were PCA ($n = 33$ solutions, 61%), ordinary least squares ($n = 10$, 19%), principal axis factoring ($n = 8$, 15%), weighted least squares ($n = 2$, 4%), and maximum likelihood ($n = 1$, 2%). Researchers generated factor analysis and PCA solutions using ordinal ($n = 49$, 91%) or nominal (binary) data ($n = 5$, 9%), and there was no evidence that polychoric correlations in these analyses. Although the researchers were principally working with ordinal data, the researchers methods reflected those recommended for continuous (interval, ratio) data. These findings are perhaps not surprising, given that guidance on data extraction methods for factor analysis with ordinal data has been limited compared to that for continuous data (e.g., Schmitt, 2011).

6.3. Recommendations for future practice

In making decisions about which data extraction methods may be most appropriate, researchers should first consider the type of measurement they have used (e.g., nominal, ordinal, interval, or ratio; Stevens, 1946). As has been highlighted in this section, the type of measurement used has clear implications for deciding which data extraction methods may be most appropriate. When analysing continuous data (i.e., interval or ratio data), the maximum likelihood method, principal axis factoring, and ordinary least squares (called unweighted least squares in SAS and SPSS) are viable options for the extraction of factors. The situations in which the maximum likelihood method has been shown to be superior to other methods are ones that are fairly common in behavioural research. For this reason, researchers should consider using the maximum likelihood method to perform the extraction of factors. The use of extraction methods other than maximum likelihood, however, may be warranted to achieve specific objectives (e.g., researchers who are concerned that weak factors have not been adequately recovered may wish to choose to use principal axis factoring or the ordinary least squares method).

In most analyses, however, researchers are likely to be working with ordinal data. When factor analysing ordinal data, researchers are encouraged to use ordinary least squares estimations with polychoric correlations. With respect to popular commercial software packages, factor analysis with polychoric correlations can be performed in SPSS with the assistance of an extension package (Basto and Pereira, 2012; Courtney, 2013) and in SAS from the Second Maintenance Release of SAS 9.3 (SAS Institute, 2013). Non-commercial software for factor analysis is also available (Lee, 2010; Lorenzo-Seva and Ferrando, 2006).

7. Deciding upon factor rotation methods

7.1. Review of literature

The rotation of factors is commonly performed with the goal of obtaining a solution that is more parsimonious and easier to interpret than the initial factor extraction (Bandalos and Boehm-Kaufman, 2009; Schmitt, 2011). Rotation assists researchers to achieve (or approximate) a simple structure (Thurstone, 1947), which involves each row (item) in an oblique factor matrix having at least one zero (also see Yates, 1987). This definition allows for complex factor patterns, whereby each item can have a non-zero loading on more than one factor. Simple structure does not mean that each item has a non-zero loading on only one factor (a situation referred to as perfect simple structure).

At the most basic level, researchers have the choice of orthogonal (when factors are uncorrelated) or oblique (when factors are correlated) forms of rotation. This choice requires researchers to be aware of the likelihood of the factors they are measuring being correlated (Schmitt, 2011). Because nursing researchers (and social scientists) often measure factors that are typically correlated (e.g., psychological and social factors), oblique rotation is likely to be appropriate for many studies. Furthermore, if the factors are orthogonal, and an oblique form of rotation is used, then a similar outcome would be achieved through oblique rotation as would have been produced using orthogonal rotation (Harman, 1976). Therefore, oblique rotations would seem to be most appropriate in nursing, and social science, research.

Evidence suggests that different rotation methods can have a substantial influence on the solutions generated (Browne, 2001; Sass and Schmitt, 2010; Schmitt and Sass, 2011). When correlation matrices represent perfect simple structures, there are minimal differences between rotation methods (Schmitt and Sass, 2011). As structures become more complex, however, substantial differences between rotation methods arise with respect to the factor pattern loadings and the correlations between factors. The presence of such differences means that researchers need to select rotation methods that are consistent with their research objectives (see Browne, 2001; Sass and Schmitt, 2010; Schmitt and Sass, 2011 for reviews of rotation methods).

One finding of note is that there is a trade-off between (a) smaller cross-loadings and larger correlations between factors, and (b) larger cross-loadings and smaller correlations between factors (Schmitt and Sass, 2011). This finding has clear practical application. In some situations (e.g., in identifying items within an instrument that may be of suspect quality, highlighting items that may load on multiple factors, and revealing highly cross-loading items that may be candidates for removal to reduce the correlations between factors) generating rotated solutions with high cross-loadings may be advantageous. Possible rotation methods that could be used to generate solutions with high cross-loadings include two forms of rotation in the Crawford–Ferguson (CF) family (Crawford and Ferguson, 1970): CF-Equamax and CF-Facparsim

(Sass and Schmitt, 2010; Schmitt and Sass, 2011). By contrast, if less complex solutions (e.g., simple structure) are expected and the goal is to reduce cross-loadings, rotation methods such as Geomin (Yates, 1987) and Direct Quartimin (Jennrich and Sampson, 1966) may be more appropriate choices. Ultimately, researchers may need to generate solutions using multiple methods and determine which one is most appropriate (Sass and Schmitt, 2010; Schmitt and Sass, 2011).

7.2. Assessment of practices evident in nursing journals

Researchers publishing in nursing journals used orthogonal rotation (Varimax, $n = 24$, 44%) more readily than oblique methods (Promax, $n = 10$, 19%; Direct Oblimin, $n = 6$, 11%; Geomin, $n = 2$, 4%; unspecified oblique methods, $n = 3$, 6%). For seven solutions (13%) no rotation was used, and for two solutions (4%) no rotation method was stated.

The common use of the Varimax method is concerning, because orthogonal methods rest on the assumption that the factors derived are uncorrelated. With respect to many of the constructs measured in nursing research (like the social sciences), using a model in which factors are unrelated does not reflect reality. The problem of expecting human traits to be orthogonal been known for decades (Thomson, 1939). Developing effective methods of oblique rotation, however, proved to be more challenging than orthogonal methods. The universal acceptance of Varimax can perhaps be attributed to its earlier development compared to oblique methods, the initial difficulties experienced with the computerisation of oblique methods, and the effectiveness with which Varimax was able to generate solutions, at a time when technological progress placed factor analysis in the hands of a much broader number of (less well trained) researchers (see Browne, 2001). There is a need for researchers publishing in the nursing literature to break from conventional practices and use oblique rotational methods more often, because they are better suited for measuring psychosocial constructs.

7.3. Recommendations for future practice

Unless they have a sound reason to do otherwise, researchers should use oblique methods of rotation. The choice of which oblique rotational method to use depends somewhat on the goals and outcomes of the analysis. If the aim is to produce solutions with smaller cross-loadings and larger correlations between factors, Geomin and Direct Quartimin are appropriate choices. Alternatively, if the intention is to generate solutions with larger cross-loadings and smaller correlations between factors, methods such as CF-Equamax and CF-Facparsim should be considered. At this point in time, there is no automated way of deciding which rotational method produces better outcomes, so human judgement with respect to what solution is best in necessary. Consistent with the exploratory purposes of this type of factor analysis, researchers may wish to generate multiple solutions, with the outcomes of each solution informing the specification of the next. Researchers may also wish to generate multiple solutions to serve different purposes.

8. Final comments

Factor analysis is a statistical method that has undergone continuous development for over 100 years, and for which evidence-based refinements to best practice recommendations are likely to continue into the foreseeable future. For the most part, however, the research published in high impact nursing journals is not reflective of the advancements in knowledge that have occurred over time. Researchers have tended to use popular rules of thumb, rather than evidence-based recommendations, when making decisions about the performance of factor analysis. The findings from the assessment show distinct preferences for (a) ratio-based heuristics for determining sample sizes, (b) PCA over factor analysis, (c) eigenvalues greater than one and scree test rules for deciding upon the numbers of factors/components to retain, and (d) the Varimax method of rotation. Without re-analysing the data in each of the included studies, we cannot be sure of the effects of these preferences on the solutions that have been published. Research has clearly shown, however, that these preferences can have adverse impacts on the validity of the solutions generated (see, for example, Fabrigar et al., 1999).

The popularity of PCA, the eigenvalues greater than one rule, and Varimax rotation dates back to Kaiser's (1958, 1960, 1970) contributions, in which he showed that this combination of methods produced meaningful outcomes for several classic data sets (Widaman, 2007). These methods were recommended at a time when limited computing power was a significant practical consideration (Jackson and Chan, 1980; Kaiser, 1960), difficulties were being experienced with developing effective computerised methods for generating solutions using oblique rotations (Browne, 2001), and substantial work had yet to occur on viable alternatives to Kaiser's preferred methods. Given the advancements in both computer technology and research on factor analysis since Kaiser's papers were published, it is unfortunate that many researchers publishing in nursing journals use heuristics developed over half a century ago, which evidence-based recommendations have now superseded.

For two main reasons, the design and availability of commercial statistical software may partially explain the poor findings in this study. First, the layout of options in some software packages has PCA, the eigenvalues greater than one rule, and Varimax rotations as prominent (and, often, default) options available for selection. The positioning of PCA within factor analysis modules may also contribute to the perpetuation of the myth that PCA is a type of factor analysis. Second, some of the methods recommended in the literature (e.g., parallel analysis, data extraction using polychoric correlations) have yet to become fully integrated in commercial software packages. As we have outlined, however, it is possible to follow the recommendations set out in this paper using commercial software.

Although this paper addresses some of the major decisions that need to be made when undertaking factor analysis, we, by no means, cover all the issues that researchers should consider. For example, we have not considered what to do about missing data or the impact of the number of response options on the validity and reliability of generated solutions. Researchers should be aware, however, that a recent simulation study found minimal differences between alternative methods of imputating missing data in the context of factor analysis, but that the expectation-maximisation method (Dempster et al., 1977) performed comparatively better than other methods (Chen et al., 2012). They should also consider evidence suggesting that the optimum number of ordinal response options is between four and seven (Lozano et al., 2008). When there are less than four categories, the validity and reliability of solutions decrease, whereas including more than seven categories results in no appreciable gains in psychometric properties. To address other questions with respect to factor analysis, researchers should consult the literature.

Our purpose of citing extensively in this paper was to give readers a sense of where some of the more popular heuristics originated and the evidence supporting, and opposing, their use. Frequently, textbooks and papers incorporating instruction on factor analysis make scant reference to the literature, which leaves readers with little ability to make judgements about whether what they are reading is informed by the latest research or repetitions of conventional wisdom and practices. In some instances, the use of out-dated heuristics has, unfortunately, been reinforced. Through our use of citations, we have provided an audit trail to the origins of ideas and research finding on factor analysis. In this way, we have sought to draw the attention of journal editors, manuscript reviewers, and researchers to the problems inherent in following popular heuristics and to recent (and, in some cases, not so recent) research on factor analysis in the hope that practices may improve.

Factor analysis is a statistical procedure that is frequently used in nursing research and other fields of scientific endeavour. The findings of the assessment, however, suggest that researchers publishing in high impact nursing journals routinely use out-dated heuristics, rather than evidence-based practices, in their decision making when performing factor analysis. We hope that this paper provides a primer for researchers and contributes to the improvement of statistical practices.

Conflict of interest

None declared.

References

Allen, I.E., Seaman, C.A., 2007. Likert scales and data analyses. Quality Progress 40 (7) 64–65.
Bandalos, D.L., Boehm-Kaufman, M.R., 2009. Four common misconceptions in exploratory factor analysis. In: Lance, C.E., Vandenberg, R.J. (Eds.), Statistical and Methodological Myths and Urban Legends: Doctrine, Verity, and Fable in the Organizational and Social Sciences. Taylor & Francis, New York, NY, pp. 61–87.
Bartholomew, D., Knott, M., Moustaki, I., 2011. Latent Variable Models and Factor Analysis: A Unified Approach. Wiley, Hoboken, NJ.
Bartlett, M.S., 1950. Tests of significance in factor analysis. British Journal of Statistical Psychology 3, 77–85.
Bartlett, M.S., 1951. A further note on tests of significance in factor analysis. British Journal of Statistical Psychology 4, 1–2.
Basto, M., Pereira, J.P., 2012. An SPSS R-Menu for ordinal factor analysis. Journal of Statistical Software 46 (4) , Retrieved from:http://www.jstatsoft.org/.
Bentler, P.M., Kano, Y., 1990. On the equivalence of factors and components. Multivariate Behavioral Research 25, 67–74.
Bernstein, I.H., Teng, G., 1989. Factoring items and factoring scales are different: spurious evidence for multidimensionality due to item categorization. Psychological Bulletin 105, 467–477.
Bock, R.D., Moustaki, I., 2007. Item response theory in a general framework. In: Rao, C.R., Sinharay, S. (Eds.), Handbook of Statistics. North-Holland, Amsterdam, Netherlands, pp. 469–514.
Boomsma, A., Hoogland, J.J., 2001. The robustness of LISREL modeling revisited. In: Cudeck, R., Du Toit, S.H.C., Sörbom, D. (Eds.), Structural Equation Models: Present and Future. A Festschrift in Honor of Karl Jöreskog. Scientific Software International, Lincolnwood, IL, pp. 139–168.
Briggs, N.E., MacCallum, R.C., 2003. Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. Multivariate Behavioral Research 38, 25–56.
Browne, M.W., 1968. A comparison of factor analytic techniques. Psychometrika 33, 267–334.
Browne, M.W., 1973. Generalized least squares estimators in the analysis of covariance structures. In: Aigner, D.J., Goldberger, A.S. (Eds.), Latent Variables in Socio-economic Models. North-Holland, Amsterdam, Netherlands, pp. 205–226.
Browne, M.W., 2001. An overview of analytic rotation in exploratory factor analysis. Multivariate Behavioral Research 36, 111–150.
Carifio, J., Perla, R.J., 2007. Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. Journal of Social Sciences 3, 106–116.
Cattell, R.B., 1966. The scree test for the number of factors. Multivariate Behavioral Research 1, 245–276.
Cattell, R.B., 1978. The Scientific use of Factor Analysis in Behavioral and Life Sciences. Plenum, New York, NY.
Chen, S.-F., Wang, S., Chen, C.-Y., 2012. A simulation study using EFA and CFA programs based the impact of missing data on test dimensionality. Expert Systems with Applications 39, 4026–4031.
Cho, S.-J., Feiming, L., Bandalos, D.L., 2009. Accuracy of the parallel analysis procedure with polychoric correlations. Educational & Psychological Measurement 69, 748–759.
Conway, J.M., Huffcutt, A.I., 2003. A review and evaluation of exploratory factor analysis practices in organizational research. Organizational Research Methods 6, 147–168.
Courtney, M.G.R., 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 (8) . Retrieved from: http://pareonline.net/
Crawford, C.B., Ferguson, G.A., 1970. A general rotation criterion and its use in orthogonal rotation. Psychometrika 35, 321–332.
Crawford, C.B., Koopman, P., 1979. Note: inter-rater reliability of scree test and mean square ratio test of number of factors. Perceptual and Motor Skills 49, 223–226.
Crawford, A.V., Green, S.B., Levy, R., Lo, W.-J., Scott, L., Svetina, D., Thompson, M.S., 2010. Evaluation of parallel analysis methods for determining the number of factors. Educational & Psychological Measurement 70, 885–901.
Croasmun, J.T., Ostrom, L., 2011. Using Likert-type scales in the social sciences. Journal of Adult Education 40 (1) 19–22.
Cudeck, R., 2007. Factor analysis in the year 2004: still spry at 100. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlbaum, Mahwah, NJ, pp. 1–7.
de Winter, J.C.F., Dodou, D., 2012. Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. Journal of Applied Statistics 39, 695–710.
de Winter, J.C.F., Dodou, D., Wieringa, P.A., 2009. Exploratory factor analysis with small sample sizes. Multivariate Behavioral Research 44, 147–181.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39, 1–38.
Fabrigar, L.R., Wegener, D.T., 2012. Exploratory Factor Analysis. Oxford University Press, New York, NY.
Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J., 1999. Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods 4, 272–299.
Fava, J.L., Velicer, W.F., 1992. An empirical comparison of factor, image, component, and scale scores. Multivariate Behavioral Research 27, 301–322.

Forero, C.G., Maydeu-Olivares, A., Gallardo-Pujol, D., 2009. Factor analysis with ordinal indicators: a Monte Carlo study comparing DWLS and ULS estimation. Structural Equation Modeling: A Multidisciplinary Journal 16, 625–641.
Fuller Jr., E.L., Hemmerle, W.J., 1966. Robustness of the maximum-likelihood estimation procedure in factor analysis. Psychometrika 31, 255–266.
Garrido, L.E., Abad, F.J., Ponsoda, V., 2012. A new look at Horn's parallel analysis with ordinal variables. Psychological Methods (Advance online publication).
Gaskin, C.J., Happell, B., 2013. Power of mental health nursing research: a statistical analysis of studies in the International Journal of Mental Health Nursing. International Journal of Mental Health Nursing 22, 69–75.
Glorfeld, L.W., 1995. An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. Educational and Psychological Measurement 55, 377–393.
Gorsuch, R.L., 1983. Factor Analysis. Erlbaum, Hillsdale, NJ.
Green, S.B., Levy, R., Thompson, M.S., Lu, M., Lo, W.-J., 2012. A proposed solution to the problem with using completely random data to assess the number of factors with parallel analysis. Educational and Psychological Measurement 72, 357–374.
Guttman, L., 1953. Image theory for the structure of quantitative variates. Psychometrika 18, 277–296.
Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2010. Multivariate Data Analysis: A Global Perspective. Pearson, Upper Saddle River, NJ.
Harman, H.H., 1960. Modern Factor Analysis. University of Chicago Press, Chicago, IL.
Harman, H.H., 1976. Modern Factor Analysis, 3rd ed. rev. University of Chicago Press, Chicago, IL.
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, 202–226.
Holgado-Tello, F., Chacón-Moscoso, S., Barbero-García, I., Vila-Abad, E., 2010. Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Quality & Quantity 44, 153–166.
Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30, 179–185.
Hubbard, R., Allen, S.J., 1987. An empirical comparison of alternative methods for principal component extraction. Journal of Business Research 15, 173–190.
IBM Corp., 2012. IBM SPSS Statistics for Windows, Version 21.0. IBM Corp., Armonk, NY.
IBM Corp., 2013. Extension Bundles from IBM SPSS. IBM Corp., Armonk, NY.
Jackson, D.N., Chan, D.W., 1980. Maximum-likelihood estimation in common factor analysis: a cautionary note. Psychological Bulletin 88, 502–508.
Jakobsson, U., 2004. Statistical presentation and analysis of ordinal data in nursing research. Scandinavian Journal of Caring Sciences 18, 437–440.
Jennrich, R.I., Sampson, P.F., 1966. Rotation for simple loadings. Psychometrika 31, 313–323.
Jöreskog, K.G., 1962. On the statistical treatment of residuals in factor analysis. Psychometrika 27, 335–354.
Jöreskog, K.G., 1969. Efficient estimation in image factor analysis. Psychometrika 34, 51–75.
Jöreskog, K.G., 1994. On the estimation of polychoric correlations and their asymptotic covariance matrix. Psychometrika 59, 381–389.
Jöreskog, K.G., 2007. Factor analysis and its extensions. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlbaum, Mahwah, NJ, pp. 47–77.
Jöreskog, K.G., Goldberger, A.S., 1972. Factor analysis by generalized least squares. Psychometrika 37, 243–260.
Jöreskog, K.G., Moustaki, I., 2001. Factor analysis of ordinal variables: a comparison of three approaches. Multivariate Behavioral Research 36, 347–387.
Kaiser, H.F., 1958. The Varimax criterion for analytic rotation in factor analysis. Psychometrika 23, 187–200.
Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurement 20, 141–151.
Kaiser, H.F., 1970. A second generation Little Jiffy. Psychometrika 35, 401–415.
Kaiser, H.F., Caffrey, J., 1965. Alpha factor analysis. Psychometrika 30, 1–14.
Katsikatsou, M., Moustaki, I., Yang-Wallentin, F., Jöreskog, K.G., 2012. Pairwise likelihood estimation for factor analysis models with ordinal data. Computational Statistics & Data Analysis 56, 4243–4258.
Kline, P., 1986. A Handbook of Test Construction: Introduction to Psychometric Design. Methuen, London, UK.
Lawley, D.N., 1940. The estimation of factor loadings by the method of maximum likelihood. Proceedings of the Royal Society of Edinburgh, Section A 50, 64–82.
Lee, S.-M., 2010. A review of CEFA Software: comprehensive exploratory factor analysis program. International Journal of Testing 10, 95–103.
Lee, C.-T., Zhang, G., Edwards, M.C., 2012. Ordinary least squares estimation of parameters in exploratory factor analysis with ordinal data. Multivariate Behavioral Research 47, 314–339.
Liu, O.L., Rijmen, F., 2008. A modified procedure for parallel analysis of ordered categorical data. Behavior Research Methods 40, 556–562.
Lorenzo-Seva, U., Ferrando, P.J., 2006. FACTOR: a computer program to fit the exploratory factor analysis model. Behavior Research Methods 38, 88–91.
Lorenzo-Seva, U., Timmerman, M.E., Kiers, H.A.L., 2011. The Hull method for selecting the number of common factors. Multivariate Behavioral Research 46, 340–364.
Lozano, L.M., García-Cueto, E., Muñiz, J., 2008. Effect of the number of response categories on the reliability and validity of rating scales. Method: European Journal of Research Methods for the Behavioral and Social Sciences 4, 73–79.
MacCallum, R.C., Tucker, L.R., 1991. Representing sources of error in the common-factor model: implications for theory and practice. Psychological Bulletin 109, 502–511.
MacCallum, R.C., Widaman, K.F., Zhang, S., Hong, S., 1999. Sample size in factor analysis. Psychological Methods 4, 84–99.
MacCallum, R.C., Widaman, K.F., Preacher, K.J., Hong, S., 2001. Sample size in factor analysis: the role of model error. Multivariate Behavioral Research 36, 611–637.
MacCallum, R.C., Browne, M.W., Cai, L., 2007. Factor analysis models as approximations. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlbaum, Mahwah, NJ, pp. 153–175.
McArdle, J.J., 1990. Principles versus principals of structural factor analyses. Multivariate Behavioral Research 25, 81–87.
Mundfrom, D.J., Shaw, D.G., Tian Lu, K., 2005. Minimum sample size recommendations for conducting factor analyses. International Journal of Testing 5, 159–168.
Neuhaus, J.O., Wrigley, C., 1954. The quartimax method: an analytic approach to orthogonal simple structure. British Journal of Statistical Psychology 7, 81–91.
Nunnally, J.C., 1978. Psychometric Theory. McGraw-Hill, New York, NY.
O'Connor, B.P., 2000. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods, Instruments, & Computers 32, 396–402.
Preacher, K.J., MacCallum, R.C., 2003. Repairing Tom Swift's electric factor analysis machine. Understanding Statistics 2 (1) 13–43.
Rao, C.R., 1955. Estimation and tests of significance in factor analysis. Psychometrika 20, 93–111.
Ruscio, J., Roche, B., 2012. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. Psychological Assessment 24, 282–292.
SAS Institute, 2013. Base SAS® 9.4 Procedures Guide: Statistical Procedures. SAS Institute, Cary, NC.
Sass, D.A., Schmitt, T.A., 2010. A comparative investigation of rotation criteria within exploratory factor analysis. Multivariate Behavioral Research 45, 73–103.
Schmitt, T.A., 2011. Current methodological considerations in exploratory and confirmatory factor analysis. Journal of Psychoeducational Assessment 29, 304–321.
Schmitt, T.A., Sass, D.A., 2011. Rotation criteria and hypothesis testing for exploratory factor analysis: implications for factor pattern loadings and interfactor correlations. Educational and Psychological Measurement 71, 95–113.
Schönemann, P.H., Steiger, J.H., 1978. On the validity of indeterminate factor scores. Bulletin of the Psychonomic Society 12, 287–290.
Snook, S.C., Gorsuch, R.L., 1989. Component analysis versus common factor analysis: a Monte Carlo study. Psychological Bulletin 106, 148–154.
Spearman, C., 1904. General intelligence, objectively determined and measured. American Journal of Psychology 15, 201–292.
Stevens, S.S., 1946. On the theory of scales of measurement. Science 103, 677–680.
Tabachnick, B.G., Fidell, L.S., 2012. Using Multivariate Statistics. Pearson, Boston, MA.
Thomson, G.H., 1934. Hotelling's method modified to give Spearman's g. Journal of Educational Psychology 25, 366–374.

Thomson, G.H., 1939. The factorial analysis of ability: I. The present position and problems confronting us. British Journal of Psychology. General Section 30, 71–77.

Thurstone, L.L., 1947. Multiple Factor Analysis. University of Chicago Press, Chicago, IL.

Velicer, W.F., 1976a. Determining the number of components from the matrix of partial correlations. Psychometrika 41, 321–327.

Velicer, W.F., 1976b. The relation between factor score estimates, image scores, and principal component scores. Educational and Psychological Measurement 36, 149–159.

Velicer, W.F., 1977. An empirical comparison of the similarity of principal component, image, and factor patterns. Multivariate Behavioral Research 12, 3–22.

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

Velicer, W.F., Jackson, D.N., 1990a. Component analysis versus common factor analysis: some further observations. Multivariate Behavioral Research 25, 97–114.

Velicer, W.F., Jackson, D.N., 1990b. Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. Multivariate Behavioral Research 25, 1–28.

Velicer, W.F., Eaton, C.A., Fava, J.L., 2000. Construct explication through factor or component analysis: a review and evaluation of alternative procedures for determining the number of factors or components. In: Goffin, R.D., Helmes, E. (Eds.), Problems and Solutions in Human Assessment: Honoring Douglas N. Jackson at Seventy. Kluwer Academic, New York, NY, pp. 41–71.

Watson, R., Thompson, D.R., 2006. Use of factor analysis in Journal of Advanced Nursing: literature review. Journal of Advanced Nursing 55, 330–341.

Widaman, K.F., 1990. Bias in pattern loadings represented by common factor analysis and component analysis. Multivariate Behavioral Research 25, 89–95.

Widaman, K.F., 1993. Common factor analysis versus principal component analysis: differential bias in representing model parameters? Multivariate Behavioral Research 28, 263–311.

Widaman, K.F., 2007. Common factors versus components: principals and principles, errors and misconceptions. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlbaum, Mahwah, NJ, pp. 177–203.

Wirth, R.J., Edwards, M.C., 2007. Item factor analysis: current approaches and future directions. Psychological Methods 12, 58–79.

Woltz, D.J., Gardner, M.K., Kircher, J.C., Burrow-Sanchez, J.J., 2012. Relationship between perceived and actual frequency represented by common rating scale labels. Psychological Assessment 24, 995–1007.

Yates, A., 1987. Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis. State University of New York, Albany, NY.

Yuan, K.-H., Marshall, L.L., Bentler, P.M., 2002. A unified approach to exploratory factor analysis with missing data, nonnormal data, and in the presence of outliers. Psychometrika 67, 95–122.

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

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