Understanding and Using Factor Scores: Considerations for the Applied Researcher

Christine DiStefano

Min Zhu

Diana Mîndrilã

Follow this and additional works at: https://scholarworks.umass.edu/pare

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Understanding and Using Factor Scores: Considerations for the Applied Researcher

Christine DiStefano
Min Zhu
Diana Mîndrilă
University of South Carolina

Following an exploratory factor analysis, factor scores may be computed and used in subsequent analyses. Factor scores are composite variables which provide information about an individual's placement on the factor(s). This article discusses popular methods to create factor scores under two different classes: refined and non-refined. Strengths and considerations of the various methods, and for using factor scores in general, are discussed.

Exploratory factor analysis (EFA) has been used as an analytical tool in educational research. The methods may be used with novel or exploratory research scenarios as a precursor to latent variable modeling or confirmatory factor analyses (CFA) (Schumaker & Lomax, 2004). However, in many research situations, EFA is used as the focal methodology. Practitioners may use EFA for a variety of purposes such as reducing a large number of items from a questionnaire or survey instrument to a smaller number of components, uncovering latent dimensions underlying a data set, or examining which items have the strongest association with a given factor. Once a researcher has used EFA and has identified the number of factors or components underlying a data set, he/she may wish to use the information about the factors in subsequent analyses (Gorsuch, 1983). For example, researchers may want to identify an individual's placement or ranking on the factor(s), use the information with hypothesis tests to determine how factor scores differ between groups, or to incorporate factor information as part of a regression or predictive analysis. To use EFA information in follow-up studies, the researcher must create scores to represent each individual's placement on the factor(s)

identified from the EFA. These *factor scores*[1] may then be used to investigate the research questions of interest.

This article will describe ways in which a researcher may create factor scores following an EFA and will discuss the advantages and disadvantages among the methods. There are two primary classes of computing factor scores: refined methods to develop factor scores require technical analyses, while non-refined methods involve non-sophisticated procedures. In this article, we discuss issues related to computing factor scores so practitioners may make informed decisions when choosing among methods.

Using and Computing Factor Scores

Creation and use of factor scores is an option in EFA and with covariance structural modeling (e.g., CFA, structural equation modeling) situations. The distinction of which methodology to use (EFA, CFA, structural

---

[1] We recognize that the EFA literature makes a distinction between factor scores and factor score estimates, where factor scores generally refer to situations where the generated factor scores are unique and factor score estimates relate to solutions where there can be more than one possible solution for the factor score. To simplify the discussion, this article will refer to factor scores meaning all types of scores indicating relative placement on an identified factor following an EFA.

Practical Assessment, Research, and Evaluation, Vol. 14 [2009], Art. 20

Practical Assessment, Research & Evaluation, Vol 14, No 20
Page 2
DiStefano, Zhu & Mîndrilă, Computing factor Scores

models) depends on many issues, such as the goal of the project, the nature of the work (i.e., exploratory or confirmatory research), and even issues such as researchers' knowledge of methodology, statistical techniques, and software. While latent variable modeling procedures are very popular, use of factor scores in the EFA framework is taught in graduate courses in research methods, included in many multivariate textbooks (e.g., Hair, Black, Babin, Anderson, & Tatham, 2006) and used in educational research situations.

To examine situations where factor scores were used, a brief literature review of peer reviewed articles from the social sciences was conducted using the PSYCINFO database. The key words "factor analysis" and "factor scores" were used to identify articles. We examined recent articles published between the years 2000 - 2009, inclusive. The search uncovered a total of 229 application articles that created and used factor scores in subsequent analyses. The articles spanned a variety of disciplines including education, psychology, public health and law. Factor scores were used for various purposes in the field of educational research. For example, Kawashima and Shiomi (2007) used EFA with a thinking disposition scale. The analyses uncovered four factors related to high school students' attitudes towards critical thinking. Using students' factor scores, Analysis of Variance was conducted by factor to investigate student differences in attitude by grade level and gender. Similarly, EFA was adopted by Bell, McCallum, and Cox (2003) in their research of cognitive elements underlying reading. After the factor solution was determined, factor scores were calculated for each factor, and were used in the follow-up multiple regression analyses to investigate the capability of the factors in predicting selected reading and writing skills.

While articles which used factor scores with EFA and also confirmatory factor analysis CFA procedures were noted in our brief literature review, the majority of these articles (123 or 53.7%) used factor scores following EFA rather than CFA procedures [2]. Additionally, many studies using factor scores did not clarify the computation procedure used to create the factor scores.

Although factor scores following EFA are still in use, the practice has been controversial in the social sciences for many years (e.g., Glass & Maguire, 1966).

For example, many factor score methods are built on the assumption that the resulting factor scores will be uncorrelated; however, orthogonal factors are often the rarity rather than the norm in educational research. Increased knowledge of the requirements underlying many of the factor score methods may provide assistance to researchers interested in using these techniques.

There are two main classes of factor score computation methods: refined and non-refined. Non-refined methods are relatively simple, cumulative procedures to provide information about individuals' placement on the factor distribution. The simplicity lends itself to some attractive features, that is, non-refined methods are both easy to compute and easy to interpret. Refined computation methods create factor scores using more sophisticated and technical approaches. They are more exact and complex than non-refined methods and provide estimates that are standardized scores.

Non-refined Methods

Non-refined methods are simple to use. Under the class of non-refined methods, various methods exist to produce factor scores. The most frequently used methods are described below.

1. Sum Scores by Factor

One of the simplest ways to estimate factor scores for each individual involves summing raw scores corresponding to all items loading on a factor (Comrey & Lee, 1992). If an item yields a negative factor loading, the raw score of the item is subtracted rather than added in the computations because the item is negatively related to the factor. For this method (as well as for the following non-refined methods) average scores could be computed to retain the scale metric, which may allow for easier interpretation. Also, average scores may be useful to foster comparisons across factors when there are differing numbers of items per factor.

The sum score method may be most desirable when the scales used to collect the original data are "untested and exploratory, with little or no evidence of reliability or validity" (Hair et al, 2006, p. 140). In addition, summed factor scores preserve the variation in the original data. Tabeachinck and Fidell (2001) noted that this approach is generally acceptable for most exploratory research situations.

While sum scores may be acceptable for many studies, there are some considerations. First, all items on

---

[2] Of the 229 application articles using factor scores: 123 articles (53.7%) used EFA, 43 articles (18.8%) used latent variable modeling techniques; and 63 articles (27.5%) did not provide sufficient information on the methodology used.

DiStefano et al.: Understanding and Using Factor Scores: Considerations for the App

Practical Assessment, Research & Evaluation, Vol 14, No 20   *Page* 3
DiStefano, Zhu & Mîndrilă, Computing factor Scores

a factor are given equal weight, regardless of the loading value. Therefore, items with relatively low loading values are given the same weight in the factor score as items with higher loading values. If items are on different metrics, ignoring different amounts of variability in the observed variables might result in less reliable factor scores. A remedy is to standardize variables on different metrics before running EFA. Lastly, summing items is straight forward if simple structure is present. The researcher must decide how to accommodate cross-loading items and how the items may impact interpretation if the items used to compute each factor score are not independent.

2. Sum Scores – Above a Cut-off Value

An easy way to consider an item's relationship to the factor when creating a factor score is to include only items with loading values above a cut-off value in the computations. By doing so, researchers are only using "marker" variables in the computation. However, the cut-off value to use is an arbitrary decision. For example, one researcher may include items above a cut-off of .30 while another researcher may include items above a higher level. While this method only includes items above the cut-off in the calculations, the variability in the raw data is not preserved.

3. Sum Scores - Standardized Variables

As compared to the previous two methods, summing standardized variables involves a somewhat more sophisticated computation approach. This option is recommended to deal with observed variables that may vary widely with respect to the standard deviation values of the raw data. Before summing, raw scores are standardized to the same mean and standard deviation. The researcher may decide to sum standardized scores of all items loaded on a factor or to decide to sum scores for items with a loading values above a cut-off value. However, this method is not necessarily more advantageous than the previous methods if standard deviations of the raw data do not vary widely.

4. Weighted Sum Scores

The preceding methods do not involve item loading values in the computations, thereby disregarding the strength (or lack of strength) for each item. As a remedy, sum scores can be created where the factor loading of each item is multiplied to the scaled score for each item before summing. This method can be applied to all the items loaded on one factor, or only to items with factor loadings above a specific cut-off value. Further, this

method can be conducted after scaling the items to the same mean and standard deviation.

Because different weights (i.e., factor loading values) are applied to each item, one advantage of the weighted sum score method is that items with the highest loadings on the factor would have the largest effect on the factor score. However, one of the potential problems with this method is that the factor loadings may not be an accurate representation of the differences among factors due to a researcher's choice of extraction model and/or rotation method. In other words, to simply weight items based on factor loadings might not result in a significant improvement over the previous methods.

Non-refined factor scores are, in general, thought to be more stable across samples than refined methods (Grice & Harris, 1998). This means that the obtained results do not heavily depend on the particular sample used. However, without a sophisticated technical computation procedure, researchers should take caution when creating and using this class of factor scores. For example, non-refined methods do not achieve a set mean and/or standard deviation for each of the factor scores. Instead, the mean and standard deviation of the factors will be dependent upon the characteristics of the items (e.g., scale of measurement, variability in data, etc.). Also, non-refined methods may produce factor scores which are correlated, even if the EFA solution is orthogonal (Glass & Maguire, 1966). While many situations involve oblique EFA solutions, the relationships among factors may not be accurately reproduced between factor scores. Finally, while non-refined methods are not obtained by a default routine with statistical software such as SPSS or SAS, the procedures can be easily programmed.

Refined Methods

Refined procedures may be applied when both principal components and common factor extraction methods are used with EFA. Resulting factor scores are linear combinations of the observed variables which consider what is shared between the item and the factor (i.e., shared variance) and what is not measured (i.e., the uniqueness or error term variance) (Gorsuch, 1983). The most common refined methods use standardized information to create factor scores, producing standardized scores similar to a Z-score metric, where values range from approximately -3.0 to +3.0. However,

instead of unit standard deviation, the exact value can vary.

Methods in this category aim to maximize validity by producing factor scores that are highly correlated with a given factor and to obtain unbiased estimates of the true factor scores. Furthermore, these methods attempt to retain the relationships between factors. In other words, when the EFA solution is orthogonal, the factor scores should be uncorrelated with other factors and when the solution is oblique, the correlations among factor scores should be the same as the correlations among factors (Gorsuch, 1983).

For the refined methods described, we provide information on how to execute these options using three popular statistical packages: SAS (version 9.12), SPSS (version 17), and R (version 2.9.0). We recognize that this presentation is a non-technical overview of the methods; the interested reader is referred to other sources for a more detailed discussion of the methodology (e.g., Gorsuch, 1983; Comrey & Lee, 1992) as well as the formulas underlying the methods (e.g., Hershberger, 2005). The formulas for all methods are also provided in the appendix.

1. Regression Scores.

Thurstone (1935) used a least squares regression approach to predict factor score(s). Regression factor scores predict the location of each individual on the factor or component. This procedure differs from the non-refined weighted sum method, in that the weighed sum non-refined procedure reflects the extent to which the factor or component estimated is manifested by each individual case; the method does not use an underlying model to predict an "optimal" factor score.

Following regression terminology, with this method, independent variables in the regression equation are the standardized observed values of the items in the estimated factors or components. These predictor variables are weighted by regression coefficients, which are obtained by multiplying the inverse of the observed variable correlation matrix by the matrix of factor loadings and, in the case of oblique factors, the factor correlation matrix. As opposed to a simple weighted sum, the least squares regression is a multivariate procedure, which takes into account not only the correlation between the factors and between factors and observed variables (via item loadings), but also the correlation among observed variables, as well as the correlation among oblique factors (for oblique EFA

solutions). The factor scores are the dependent variables in the regression equation.

Under this process, the computed factor scores are standardized to a mean of zero; however, the standard deviation of the distribution of factor scores (by factor) will be 1 if principal components methods are used and will be the squared multiple correlation between factors and variables (typically used as the communality estimate) if principal axis methods are used (Tabachnick & Fidell, 2001).

Regression based factor scores can be easily obtained with popular statistical software packages. SAS computes regression factor scores and the results can be saved in an output data set specified in the OUT= option of the PROC FACTOR statement. In SPSS, regression factor scores are obtained by clicking the Scores button in the Factor Analysis window, checking the "Save as variables" box in the Factor Analysis: Factor Scores window and selecting "Regression" (default) from the three options provided. In the computer package, R, regression factor scores are produced when an option scores = "regression" is added to the FACTANAL function, e.g. fact1<-FACTANAL(x, factors, scores= c("regression"), rotation= "varimax").

2. Bartlett Scores.

With Bartlett's approach, only the shared (i.e., common) factors have an impact on factor scores. The sum of squared components for the "error" factors (i.e., unique factors) across the set of variables is minimized, and resulting factor scores are highly correlated to their corresponding factor and not with other factors. However, the estimated factor scores between different factors may still correlate.

Bartlett factor scores are computed by multiplying the row vector of observed variables, by the inverse of the diagonal matrix of variances of the unique factor scores, and the factor pattern matrix of loadings. Resulting values are then multiplied by the inverse of the matrix product of the matrices of factor loadings and the inverse of the diagonal matrix of variances of the unique factor scores.

One advantage of Bartlett factor scores over the other two refined methods presented here is that this procedure produces unbiased estimates of the true factor scores (Hershberger, 2005). This is because Bartlett scores are produced by using maximum likelihood estimates – a statistical procedure which

produces estimates that are the most likely to represent the "true" factor scores.

Bartlett factor scores can be computed in SPSS under the FACTOR menu by checking the "Save as variables" box in the Factor Analysis: Factor Scores window and selecting "Bartlett" from the options provided. In R, Bartlett's factor scores can be computed using a procedure similar to that used to obtain the regression factor scores, but including the option scores= "Bartlett".

3. Anderson-Rubin Scores.

The method proposed by Anderson and Rubin (1956) is a variation of the Bartlett procedure, in which the least squares formula is adjusted to produce factor scores that are not only uncorrelated with other factors, but also uncorrelated with each other. Computation procedures are more complex than the Bartlett method and consist of multiplying the vector of observed variables by the inverse of a diagonal matrix of the variances of the unique factor scores, and the factor pattern matrix of loadings for the observed variables. Results are then multiplied by the inversion of the symmetric square root of the matrix product obtained by multiplying the matrices of eigenvectors (characteristic vectors of the matrix) and eigenvalues (characteristic roots of the matrix)[3]. The resulting factor scores are orthogonal, with a mean of 0 and a standard deviation of 1. They can be automatically generated in SPSS by selecting the Anderson and Rubin option in the Factor Analysis: Factor Scores window.

Each of the three refined methods has advantages as well as drawbacks. The main advantage of the regression method is that it maximizes validity. This means that the procedure provides the highest correlations between a factor score and the corresponding factor. Nevertheless, regression estimates are not unbiased estimates of true factor scores and could correlate with other factors and factor scores, even

when the EFA solution is orthogonal. Bartlett factor scores are also highly correlated with the factor being estimated (Gorsuch, 1983). This method has the additional advantage that factor scores only correlate with their own factor in an orthogonal solution. Finally, resulting coefficients are unbiased and, therefore, more accurate reflections of the cases' location on the latent continuum in the population. The most important disadvantage of the Bartlett approach is that there may be a relationship among the factor scores from different factors in an orthogonal solution. The Anderson-Rubin method produces factor scores that are orthogonal when the solution is orthogonal. On the other hand, even though the factor scores have reasonably high correlations with the corresponding factor, they may also correlate with other factors in an orthogonal solution, and the factor scores are not unbiased (Gorsuch, 1983). In conclusion, none of the refined methods can concomitantly maximize validity, maximize the correlation between factor scores and their parent factor, and provide uncorrelated estimates for orthogonal factors. Table 1 summarizes advantages and disadvantages of the different refined methods.

Considerations

While factor scores following EFA are relatively easy to create and may be useful for follow-up analyses, caveats to using these scores, regardless of the method used to compute them, should be noted. First, factor scores are sensitive to the factor extraction method and rotation method used to create the EFA solution. Just as researchers are likely to obtain different solutions when different extraction and/or rotation method are used, factor scores obtained with different EFA selections may be different as well. This could affect follow-up tests if factor scores for the same case differ widely across different EFA methods. Similarly, the purpose of the initial EFA should be conducive to the use of the factor scores in further analyses. Zuccaro (2007) discussed potential interpretation problems that may arise when refined methods were used to produce standardized factor scores, but other variables in the follow-up analyses were not standardized. Additionally, if EFA is used improperly (e.g., researchers are extracting components for data reduction purposes but then treating the components as latent variables), misleading hypothesis test results may occur (Zuccaro, 2007). We remind researchers to first determine if EFA is acceptable and then to using factor scores, given that the EFA met the needs of the initial research question.

---

[3] Characteristic roots (eigenvalues) and characteristic vectors (eigenvectors) are used in factor analysis for matrix decomposition. Eigenvalues are scalars (i.e., one number) that show the proportion of variance accounted for by each factor. The matrix used in the calculations is of order $m \times m$ (where m = number of factors) with eigenvalues on the diagonal and 0's on the off diagonal. Eigenvectors are vectors, which contain one value for each variable in the factor analysis. When eigenvectors are multiplied by the square root of the eigenvalue, the factor loading is produced. See Gorsuch (1983) for more information.

DiStefano, Zhu & Mîndrilă, Computing factor Scores

Table 1: Advantages and Disadvantages of Refined Methods to Compute Factor Scores

Factor scores are…	…correlated to the estimated factor. (Validity)	…uncorrelated to other orthogonal factors. (Univocality)	…uncorrelated to factor scores from other orthogonal factors. (Correlational Accuracy)	…unbiased estimates of factor score parameters.
Regression scores	Maximal	No	No	No
Bartlett factor scores	High	Yes	No	Yes
Anderson & Rubin factor scores	Acceptable	No	Yes	No

A second, and paramount, consideration when creating factor scores using refined methods is the problem of "indeterminacy" of the scores (see Gorsuch, 1983 and Grice, 2001 for detailed explanations). Indeterminacy arises from the fact that, under the common factor model, the parameters are not uniquely defined, due to the researcher's choice of the communality estimate. This means that there is not a unique solution for the factor analysis results and, theoretically, an infinite number of solutions could account for the relationships between the items and factor(s). Therefore, it also follows that the factor scores are not uniquely defined (Grice, 2001). As noted by EFA researchers, the problem of indeterminacy arises with most factor extraction techniques found under the common factor model. EFA methods which have a unique solution (i.e., determinant), such as principal component analysis and image common factor analysis, and their resulting factor scores are thought to be unique. Researchers interested in using factor scores need to be aware of the problem of indeterminacy because it could impact not only the factor scores but also the validity of decisions that rely upon these scores (Grice, 2001). For example, under some conditions, rankings of cases in a data set may vary widely based on different methods to compute factor scores, leaving a researcher unsure as to which ranking to trust.

Grice (2001) suggests researchers examine the degree of indeterminacy in their factor solutions using three different measures: (1) validity – evidence of correlational relationships between factor scores and factor(s); (2) univocality- the extent to which factor scores are adequately or insufficiently correlated with other factors in the same analysis; and (3) correlational accuracy – which reports the extent to which correlations among the estimated factor scores match the correlations among the factors themselves. Table 1 includes these three measures to illustrate differences among the refined methods. As noted by Grice, a high degree of indeterminacy may suggest that a researcher re-examine the number of factors needed or to disregard (or at least, use cautiously), scores from factors which illustrate questionable results. Interested readers should refer to Grice (2001) for an in-depth discussion of indeterminacy, illustrations of how to evaluate if it is present, and a link to SAS programs to examine solutions for indeterminacy. While popular software programs do not yet routinely provide all of these tests, a test for validity using the multiple correlation value is routinely available under the regression method of SPSS when orthogonal factors are requested and with SAS for both orthogonal and oblique solutions. Higher values of the multiple correlation suggest greater validity evidence, meaning greater determinacy of the factor scores. This information is very important for researchers to consider because EFA is an internally driven procedure, and thus, results may be sample specific (Costello & Osborne, 2005). Given the problems with EFA, researchers creating factor scores are urged to replicate the factor structure to ensure that the solution is stable before creating and using factor scores. Factor scores may also be examined for indeterminacy using the procedures described above.

A third issue deals with data quality. Once factor scores are obtained, this set of data requires screening and examination to ensure that distribution(s) of factor scores meet assumptions required by the statistical methodology to be used for follow-up testing. While recommendations for data screening and checking assumptions are common before beginning statistical

analyses, it deserves repeating in the context of factor scores. Factor scores are "new" data for a follow-up analysis and are subject to the same screening recommendations. Factor scores may be skewed and/or non-normal, especially if non-refined methods were used to create the scores. Further action (e.g., transformations) may be needed before using factor scores in subsequent analyses. Failure to properly screen the factor scores may result in results of hypothesis tests that could provide misleading or even incorrect information.

Lastly, we wish to recognize that factor scores can be computed in the context of CFA. CFA and its uses differ from EFA; however, the factor scores created through CFA are similar in the sense that they can be used to identify ranking on a latent variable and used in follow-up analyses. CFA methods have additional advantages over EFA, including conducting measurement at the latent level, distinguishing the error component from what is shared with a factor, including multiple fit indices, and allowing for much greater flexibility in constructing a model (Bollen, 1989). Factor scores computed in the CFA context typically use similar methods as described here. The interested reader should refer to discussions of factor scores in the CFA framework (e.g., Bollen, 1989).

In summary, this discussion introduced the topic of factor scores within the EFA framework, described different methods to create factor scores, and provided advantages and disadvantages among the different methods. The appendices provided summarize the preceding discussion of factor scores in a table for use as a reference tool and also provide additional information on the computation of factor scores under the refined methods. We hope that this discussion will help to illustrate choices, considerations, and caveats when creating and using factor scores.

References

Anderson, R. D., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, 5*, 111-150.

Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology, 28,* 97-104.

Bell, S. M., McCallum, R. S., & Cox, E. A. (2003). Toward a research-based assessment of dyslexia: Using cognitive measures to identify reading disabilities. *Journal of Learning Disabilities, 36*(6), 505-516.

Bollen, K. A. (1989). *Structural models with latent variables.* New York: Wiley.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillside, NJ: Erlbaum.

Costello, A. B. & Osboren, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7). Available online: http://pareonline.net/pdf/v10n7.pdf.

Glass, G. V., & Maguire, T. O. (1966). Abuses of factor scores. *American Educational Research Journal, 3*(4), 297-304.

Gorsuch, R. (1983). *Factor analysis.* Hillsdale, NJ: L. Erlbaum Associates.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*(4), 430-450.

Grice, J. W. & Harris, R. J. (1998). A comparison of regression and loading weights for the computation of factor scores. *Multivariate Behavioral Research, 33*(2), 221-247.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006), *Multivariate Data Analysis (6th ed.).* Upper Saddle River, N.J.: Pearson Education Inc.

Hershberger, S. L. (2005). Factor scores. In B. S. Everitt and D. C. Howell (Eds.) *Encyclopedia of Statistics in Behavioral Science.* (pp. 636-644). New York: John Wiley.

Kawashima, N., & Shiomi, K. (2007). Factors of the thinking disposition of Japanese high school students. *Social Behavior and Personality, 35*(2), 187-194.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics (4th Ed.).* Needham Heights, MA: Allyn & Bacon.

Thurstone, L. L. (1935). *The vectors of mind.* Chicago: University of Chicago Press. (pp. 226-231).

Zuccaro, C. (2007). Statistical alchemy--The use of factor scores in linear regression. Retrieved December 4, 2008, from http://www.esg.uqam.ca/recherche/document/2007/01-2007.pdf

Appendix 1: Non-Refined Methods to Construct Factor Scores

Method	Procedure	Advantages	Considerations
Sum Scores by Factor Sum Scores Above a Cut-off Value	Sum raw scores corresponding to all items loading on the factor. (Items with negative loadings are subtracted in the score creation.) Sometimes a cutoff loading value is used and items above the cutoff are summed.	In the metric of what is studied. Can be averaged to reflect the scale of the items. Easy to calculate and interpret. If factor scores are used in later analyses, sum scores preserve variation in the data.	Gives items equal weight when the weight of item to factor (loading values) may be very different. Cutoff is arbitrary. A higher cutoff may result in including fewer variables used, a lower cutoff will include variables with a weaker relationship to the factor.
Sum Scores - Standardized Variables	Scale raw scores to same mean and standard deviation before summing. Can apply a cutoff loading value and only add items above the cutoff.	Useful to deal with observed variables that vary widely in terms of standard deviation units. Refinement worth effort unless observed variables are reasonably similar in the size of standard deviations.	If standard deviations of raw scores are similar, sum scores without standardizing are easier to compute. No weighting given to items with higher loadings.
Weighted Sum Scores	Take into consideration the loading values in the factor score creation. Multiply the factor loading to the scale score then sum. Can be applied to items above a certain loading value or all items on a factor.	Recognizes the strength (or lack of strength) for items. Items with highest loadings have the most affect on the factor scores.	Possibility that differences in factor loadings are due to EFA extraction and rotation choices. If differences are due to EFA procedures, this method may not be better than creating summed scale scores.

Appendix 1 (continued): Refined Methods to Construct Factor Scores

Method	Procedure	Advantages	Considerations
Regression Scores	Multiple regression used to estimate (predict) factor scores. Default procedure to compute factor scores in SAS and SPSS packages; also available in R.	Factor scores are standard scores with a Mean =0, Variance = squared multiple correlation (SMC) between items and factor. Procedure maximizes validity of estimates.	Factor scores are neither univocal nor unbiased. The scores may be correlated even when factors are orthogonal.
Bartlett	Method of producing factor scores is similar to regression method, but produces estimates that are most likely to represent the true factor scores. Can be computed using SPSS or R statistical packages.	Factor scores are standard scores (Mean =0, Variance = SMC) Produces unbiased estimates. In an orthogonal solution, factor scores are not correlated with other factors (univocality). Procedure produces high validity estimates between the factor scores and factor.	The scores may be correlated even when factors are orthogonal.
Anderson-Rubin	Method of producing factor scores is similar to Bartlett, but allows factor scores to be uncorrelated when factors are orthogonal. Can be computed using SPSS.	Factor scores have a mean of 0, have a standard deviation of 1. When the factors are orthogonal, factor scores are uncorrelated as well (correlational accuracy). Factor scores have reasonably high correlations with their estimated factor (validity).	Factor scores may be correlated with the other orthogonal factors (i.e.,. not univocal). Factor scores are not unbiased.

Practical Assessment, Research & Evaluation, Vol 14, No 20   Page 10
DiStefano, Zhu & Mîndrilă, Computing factor Scores

Appendix 2: Computation Procedures for Refined Factor Scores

Factor Scores	Formulae	Where…
Regression Scores	$\hat{F}_{1xm} = Z_{1xn} B_{nxm}$	n – number of observed variables m – number of factors
Orthogonal Factors:	$B_{nxm} = R^{-1}_{nxm} A_{nxm}$	$\hat{F}$ – the row vector of m estimated factor scores
Oblique Factors:	$B_{nxm} = R^{-1}_{nxm} A_{nxm} \Phi_{mxm}$	Z – the row vector of n standardized observed variables
		B – the matrix of regression of weights for the m factors on the n observed variables
		R^{-1} – the inverse of the matrix of correlations between the n observed variables
		A and A' – pattern matrices of loadings of n observed variables on m factors or components
		Φ – the correlation matrix of the m factors
Bartlett	$\hat{F}_{1xm} = Z_{1xn} U^{-2}_{nxn} A_{nxm} (A'_{mxn} U^{-2}_{nxn} A_{nxm})^{-1}$	U^{-2} – the inverse of a diagonal matrix of the variances of the n unique factor scores
Anderson-Rubin	$\hat{F}_{1xm} = Z_{1xn} U^{-2}_{nxn} A_{nxm} G^{-1/2}$ $G_{nxn} = X_{nxn} \Lambda_{Dnxn} X'_{nxn}$	X and X' – matrices of n x n eigenvectors Λ_D – the n x n matrix of eigenvalues G - the matrix product of eigenvalues and eigenvectors G$^{-1/2}$ – the inverse of the symmetric square root of G

Formulas for refined factor methods were taken from Hershberger, S. L. (2005).

DiStefano, Zhu & Mîndrilă, Computing factor Scores

Citation

Corresponding Author

Christine DiStefano, Associate Professor
Department of Educational Studies
137 Wardlaw Hall
Columbia, SC 29208
email: distefan [at] mailbox.sc.edu