The Use of Exploratory Factor Analysis and Principal Components Analysis in Communication Research

HEE SUN PARK
Michigan State University
RENÉ DAILEY
DAISY LEMUS
University of California, Santa Barbara

Exploratory factor analysis is a popular statistical technique used in communication research. Although exploratory factor analysis (EFA) and principal components analysis (PCA) are different techniques, PCA is often employed incorrectly to reveal latent constructs (i.e., factors) of observed variables, which is the purpose of EFA. PCA is more appropriate for reducing measured variables into a smaller set of variables (i.e., components) by keeping as much variance as possible out of the total variance in the measured variables. Furthermore, the popular use of varimax rotation raises some concerns about the relationships among the factors that researchers claim to discover. This paper discusses the distinct purposes of PCA and EFA, using two data sets as examples to highlight the differences in results between these procedures, and also reviews the use of each technique in three major communication journals: Communication Monographs, Human Communication Research, *and* Communication Research.

Looking back at the last decade of research in communication, it is evident that factor analysis is a widely used statistical procedure. Because factor analysis is so popular in communication research, it is important to determine if this procedure is being utilized correctly. In 1979, McCroskey and Young expressed their concerns about the abuse or misuse of factor analysis in communication research and provided advice on how to enhance the use of factor analysis. More recently, Fabrigar, Wegener, MacCallum, and Strahan (1999) alerted psychologists about their questionable use of exploratory factor analysis as well as their inappropriate use of factor extraction and rotation procedures for their research goals. This article has a similar purpose—to examine how researchers in

Hee Sun Park (M.A., University of Hawai'i, 1998) is a visiting instructor in the Department of Communication at Michigan State University. *René Dailey* (M.A., University of Wyoming, 1998) and *Daisy Lemus* (B.A., University of Southern California, 2000) are graduate students in the Department of Communication at the University of California, Santa Barbara, CA. Please direct correspondence to the first author; email: heesun@msu.edu.

Human Communication Research, Vol. 28 No. 4, October 2002 562–577

communication conduct factor analyses in addition to showing how different decisions regarding factor extraction and rotation methods can result in different conclusions. An examination of three major communication journals (*Human Communication Research*, *Communication Monographs*, and *Communication Research*) from 1990 to 2000 leads us to conclude that researchers should use more caution when making decisions regarding factor analysis procedures for their projects.

The five major methodological issues for conducting exploratory factor analysis include (a) decisions on the appropriateness of factor analysis for the research goals, (b) the size and type of sample suitable for the variables in the study, (c) the specific procedure to fit the model to data, (d) the number of factors to be retained, and (e) the right type of rotation for interpretation (Fabrigar et al., 1999). For specific and detailed discussion of each methodological issue, excellent sources exist, such as Fabrigar et al., Wegener and Fabrigar (2000), and Gorsuch (1983). Our goal is not to provide a review on all the issues involving factor analysis, but to primarily focus on the factor extraction and rotation methods commonly used in communication research. First, we discuss the methodological differences between exploratory factor analysis (EFA) and principal components analysis (PCA) as well as differences between orthogonal and oblique rotations. Second, we illustrate the differences between EFA and PCA with two data sets. Finally, we examine how EFA and PCA are used in communication research and discuss the implications of these different uses.

EXPLORATORY FACTOR ANALYSIS VERSUS PRINCIPAL COMPONENTS ANALYSIS

The primary reason we want to discuss differences between EFA and PCA is that PCA is often mistakenly used when EFA is more appropriate. The goal of EFA is to find a latent structure of observed variables by uncovering common factors that influence the measured variables, whereas the goal of PCA is to reduce the measured variables to a smaller set of composite components that capture as much information as possible in the measured variables with as few components as possible. This difference is reflected in the way the two techniques analyze the variation in the measured variables. EFA focuses on the shared variance among the variables by separating common variance from unique variance. Common variance is what a variable shares with the other variables, while unique variance is the variance specific to each variable. In contrast, PCA does not distinguish between common and unique variances because PCA focuses on the total variation among the variables.

In order to make the methodological distinction between EFA (common factor analysis) and PCA clearer, algebraic equations of each are shown below, as presented in Dillon and Goldstein (1984).

Algebraic representation of PCA for m ($\leq p$) principal components is as follows:

$$
\begin{aligned}
PC_{(1)} &= w_{(1)1}X_1 + w_{(1)2}X_2 + \ldots\ldots + w_{(1)p}X_p \\
PC_{(2)} &= w_{(2)1}X_1 + w_{(2)2}X_2 + \ldots\ldots + w_{(2)p}X_p \\
&\vdots \\
PC_{(m)} &= w_{(m)1}X_1 + w_{(m)2}X_2 + \ldots\ldots + w_{(m)p}X_p
\end{aligned}
$$

m: the number of principal components
p: the number of measured variables
X: measured variable
PC: principal component
$w_{(i)j}$: the weight chosen for the jth measured variable to maximize the ratio of the variance of $PC_{(i)}$ to the total variation, $i = 1, 2, \ldots, m, j = 1, 2, \ldots, p$

Algebraic representation of EFA for m (< p) common factors is as follows:

$$
\begin{aligned}
X_1 &= v_{1(1)}CF_{(1)} + v_{1(2)}CF_{(2)} + \ldots\ldots + v_{1(m)}CF_{(m)} + e_1 \\
X_2 &= v_{2(1)}CF_{(1)} + v_{2(2)}CF_{(2)} + \ldots\ldots + v_{2(m)}CF_{(m)} + e_2 \\
&\vdots \\
X_p &= v_{p(1)}CF_{(1)} + v_{p(2)}CF_{(2)} + \ldots\ldots + v_{p(m)}CF_{(m)} + e_p
\end{aligned}
$$

CF: common factor
$v_{j(i)}$: the weight of the ith common factor associated with the jth measured variable, $i = 1, 2, \ldots, m, j = 1, 2, \ldots, p$
e_j: unique factor, $j = 1, 2, \ldots, p$.

The equations above clearly show the two major differences between PCA and EFA. First, EFA estimates errors (i.e., unique variance), while PCA does not. This indicates that PCA assumes that the measurement is without error. Second, in EFA, the measured variables are a function of factors; in PCA, components are a function of the measured variables. EFA tries to explain the correlations among the variables quantitatively and qualitatively, while PCA attempts to account for the variation in the variables in a way that will retain as much information as possible from the original variables (Fabrigar et al., 1999). For example, if a researcher wants to reduce measures such as income, education, and level of debt into a composite component such as socioeconomic status, PCA would be appropriate. On the other hand, if a researcher wants to examine a latent factor such as verbal aggressiveness through various variables measuring, for example, tendency to lose temper, use of insults, offensive comments, intense language, and so on, EFA would be appropriate.

PCA and EFA can produce similar outcomes, especially when the error is close to zero (Velicer, Peacock, & Jackson, 1982). However, PCA and EFA are more likely to generate different outcomes when the ratio of the number of factors to the number of measured variables is low (e.g., three variables for a factor) and when the communalities in EFA are also low (e.g., 0.40; Widaman, 1993). If researchers perform PCA in order to find latent constructs among the measured variables, they cannot estimate unique variance (i.e., error) in each of the variables. Unless an EFA is also conducted, it is unknown whether this decision was appropriate. Thus, researchers need to be aware of the different characteristics of PCA and EFA in order to make the most appropriate decisions regarding choice of procedures suitable for their research goals.

THE NUMBER OF FACTORS TO RETAIN

Regardless of whether EFA or PCA is conducted, researchers need to decide on the number of factors/components to retain. A few popular factor/component-determining methods are eigenvalues > 1 and the scree test. But these methods have been known for their inadequacy (see Fabrigar et al., 1999; Zwick & Velicer, 1982, 1986). One rationale for using eigenvalues > 1 is that the variance of a factor/component should be greater than the variance of one measured variable. However, this criterion of retaining the factors/components whose eigenvalues are greater than 1 has a tendency to overestimate the number of factors/components (Browne, 1968; Linn, 1968; Revelle & Rocklin, 1979; Zwick & Velicer, 1982) and in some cases to underestimate them (Humphreys, 1964). With the scree test, a test that plots the eigenvalues of the factors/components in descending order, the factor/component number decisions are made when the lines connecting the eigenvalues indicate a break point (i.e., a noticeable drop or the point of starting a relatively flat straight line). The scree test can be accurate especially when samples are large and factors/components are strong (Cliff & Hamburger, 1967; Linn, 1968; Zwick & Velicer, 1982). However, problems can arise when there are multiple breaks or no discernable break points.

Maximum likelihood (ML) factor extraction method provides another way to determine the number of factors. An advantage of the ML extraction method is that it provides significance testing for the number of factors. Thus, researchers can compare models with different numbers of factors and make a decision based on various fit indexes that ML estimation generates. For example, the root mean square error of approximation (RMSEA) fit index (Browne & Cudeck, 1993) and the expected cross-validation index (ECVI; Browne & Cudeck, 1989) can serve as the indicators used to determine the best number of factors. RMSEA estimates the discrepancy between the model and the data per degree of freedom for the model. RMSEA values less than .05 indicate good fit, values from .05 to .08 acceptable/reasonable fit, values from .08 to .10 marginal fit, and values above .10 unacceptable/poor fit (Browne & Cudeck, 1993). On the other hand, ECVI represents how well the model will perform with other samples. ECVI values are meaningful only when comparing more than two models with different number of factors. The smaller ECVI value, the more generalizable the model.

FACTOR ROTATION METHODS

For interpretability of factor loadings, factors are rotated in n (i.e., number of factors) dimensional space in a way to produce simple structures. The two ways to rotate factors are oblique and orthogonal. An orthogonal rotation method (e.g., varimax, equimax, quartimax, etc.) constrains factors to be independent of each other, while an oblique rotation method (e.g., promax, oblimin, quartimin, etc.) allows factors to be correlated. It is often believed that an orthogonal rotation produces a simpler and more easily interpretable structure of factors. However, this common belief (or convention of preferring varimax rotation) is unwarranted and unrealistic. The solution of an orthogonal rotation is in fact no simpler than the solution of an oblique rotation and can be misleading with the presence of significant correlations among factors (Fabrigar et al., 1999). Furthermore, many constructs in communication research cannot be expected to be independent of each other and, even if the factors are indeed unrelated, an oblique rotation will show correlations close to zero.

Overall, being more informed regarding the differences between EFA and PCA, rotation methods, and decision rules for retaining factors could facilitate more accurate choices on the part of researchers in their factor analytic projects.

ILLUSTRATION OF EFA AND PCA

In order to see how EFA and PCA can lead to different conclusions about measured variables, we used two data sets and conducted an EFA and a PCA on each. Table 1 shows loadings of items from an embarrassibility scale[1] and Table 2 shows loadings of items from an inter-

TABLE 1
Loadings for Common Factors and Principal Components Using Varimax (Orthogonal) and Promax (Oblique) Rotations: Embarrassibility

Items	*Maximum likelihood factors with varimax*			*Principle components factors with varimax*			*Maximum likelihood factors with promax*			*Principal components factors with promax*		
	1	2	3	*1*	2	3	*1*	2	3	*1*	2	3
1	**0.855**	0.045	0.045	**0.782**	0.123	0.009	**0.933**	-0.175	-0.007	**0.819**	-0.023	-0.087
2	**0.640**	0.116	-0.008	**0.676**	0.140	-0.021	**0.675**	-0.025	-0.071	**0.704**	0.022	-0.113
3	**0.495**	0.125	0.089	**0.644**	0.072	0.070	**0.505**	0.005	0.038	**0.674**	-0.061	0.001
4	**0.419**	**0.310**	0.124	**0.449**	**0.338**	0.129	**0.361**	0.239	0.034	**0.402**	0.266	0.031
5	**0.388**	0.184	0.054	**0.565**	0.132	0.026	**0.371**	0.109	-0.008	**0.581**	0.029	-0.051
6	**0.375**	**0.342**	0.072	**0.309**	**0.473**	0.046	**0.305**	0.298	-0.028	0.232	**0.457**	-0.071
7	**0.329**	**0.411**	0.179	**0.354**	**0.364**	0.284	0.226	**0.373**	0.072	0.277	0.291	0.199
8	0.250	**0.383**	0.119	**0.354**	**0.324**	0.187	0.150	**0.371**	0.019	0.296	0.261	0.103
9	0.261	**0.485**	0.178	0.228	**0.462**	**0.307**	0.128	**0.479**	0.055	0.119	**0.422**	0.216
10	0.223	**0.384**	0.126	**0.366**	**0.333**	0.148	0.120	**0.378**	0.026	**0.311**	0.275	0.058
11	0.196	**0.486**	0.089	0.262	**0.451**	0.161	0.060	**0.513**	-0.037	0.173	**0.426**	0.059
12	0.165	**0.491**	0.077	0.153	**0.598**	0.049	0.025	**0.529**	-0.049	0.037	**0.628**	-0.077
13	0.146	**0.477**	0.127	0.104	**0.589**	0.101	0.005	**0.508**	0.009	-0.021	**0.620**	-0.016
14	-0.003	**0.436**	0.021	-0.101	**0.587**	0.002	-0.140	**0.517**	-0.086	-0.230	**0.674**	-0.102

TABLE 1 Continued
Loadings for Common Factors and Principal Components Using Varimax (Orthogonal) and Promax (Oblique) Rotations: Embarrassibility

Items	*Maximum likelihood factors with varimax*			*Principle components factors with varimax*			*Maximum likelihood factors with promax*			*Principal components factors with promax*		
	1	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
15	0.110	**0.430**	-0.053	0.090	**0.481**	-0.005	-0.008	**0.495**	-0.171	-0.001	**0.519**	-0.107
16	0.088	**0.457**	0.254	0.020	**0.482**	**0.360**	-0.061	**0.476**	0.153	-0.116	**0.478**	0.288
17	0.068	**0.460**	0.223	-0.024	**0.531**	0.299	-0.082	**0.490**	0.120	-0.168	**0.551**	0.218
18	0.066	**0.382**	0.347	0.052	**0.376**	**0.441**	-0.067	**0.375**	0.272	-0.068	**0.342**	**0.391**
19	0.060	0.293	**0.572**	0.044	0.177	**0.751**	-0.061	0.228	**0.536**	-0.070	0.075	**0.760**
20	0.070	-0.016	**0.827**	0.062	-0.036	**0.786**	0.030	-0.191	**0.886**	-0.009	-0.172	**0.837**
21	0.030	0.044	**0.797**	0.015	-0.001	**0.801**	-0.031	-0.103	**0.842**	-0.069	-0.125	**0.850**
22	0.125	0.174	**0.377**	0.165	0.075	**0.520**	0.060	0.105	**0.354**	0.107	-0.028	**0.524**
23	0.205	0.241	0.018	0.176	**0.400**	-0.081	0.151	0.232	-0.052	0.116	**0.422**	-0.180
24	0.297	0.075	0.273	**0.513**	-0.118	**0.362**	0.289	-0.038	0.258	**0.540**	-0.289	**0.360**
25	0.230	0.251	0.090	0.192	**0.406**	0.006	0.171	0.224	0.022	0.123	**0.413**	-0.090
26	0.210	0.202	0.181	0.214	0.213	0.233	0.159	0.153	0.133	0.163	0.159	0.187

NOTE: Loadings in bold are values greater than 0.30. Correlations among factors were shown by 1 and 2 (0.500); 1 and 3 (0.238); as well as 2 and 3 (0.419). Correlations among components were shown by 1 and 2 (0.406); 1 and 3 (0.272); as well as 2 and 3 (0.363).

dependent self-construal scale.[2] The purpose of these illustrations is not to test the validity of each scale, but to demonstrate how EFA and PCA can yield different results.

With data on embarrassibility (see Table 1), EFA with the ML method shows an RMSEA value of .058 for a 3-factor solution and an ECVI value of 3.356. This indicates better fit than 1-factor (RMSEA = 0.085, ECVI = 4.355), 2-factor (RMSEA = 0.070, ECVI = 3.701), or 4-factor models (RMSEA = 0.062, ECVI = 3.552). Models with more than four factors yielded slightly better fit (smaller RMSEA values than 3-factor model) but show little or no improvement in ECVI values. Thus, for simplicity and interpretability of the factors,[3] the 3-factor solution appears acceptable. However, when using the criterion of selecting factors based on eigenvalues greater than 1, the PCA yields an 8-factor solution.

To compare orthogonal (i.e., varimax) and oblique (i.e., promax) rotation methods, we set both the EFA with ML extraction and the PCA to yield 3-factor solutions. The results of these analyses are provided in Table 1 and will be discussed in conjunction with the results of the self-construal analysis.

Similar procedures were used to determine the number of factors and components for the analyses regarding the self-construal data. An EFA with the ML method shows marginal fit with 3-factor (RMSEA = 0.086, ECVI = 1.334), and acceptable fit with both 4-factor (RMSEA = 0.064, ECVI = 1.188), and 5-factor (RMSEA = 0.052, ECVI = 1.173) models.[4] Thus, the 5-factor solution appears most acceptable. However, a PCA using the criteria of eigenvalues greater than 1 yields a 4-factor solution. In order to compare the results of EFA and PCA using both orthogonal and oblique rotations, the 4-factor solutions were chosen. Results are presented in Table 2.

Two patterns emerge when examining Tables 1 and 2. First, when the loadings above 0.30 are taken for each analysis and compared across EFA with ML and PCA, the majority of loadings are higher for PCA than for EFA with ML. In general, PCA loadings are lower than EFA loadings when EFA loadings are above 0.79 and PCA loadings are above 0.30 on more than one component. This observation that PCA inflates loadings when compared to EFA is consistent with the findings of Fabrigar et al. (1999) and Widaman (1993). An inflation of loadings can create problems when deciding which items to include in the factors/components and ultimately bias the interpretation of the nature of the constructs.

The second pattern apparent in the tables is that correlations among factors/components tend to be lower for PCA than for EFA with ML. Because PCA includes errors in the components (while EFA with ML does not), it underestimates relations among the components (Fabrigar et al., 1999; Widaman, 1993). With the exception of the correlation between the first and the third factors of the embarrassibility scale, the tables show

TABLE 2
Loadings for Common Factors and Principal Components Using Varimax and Promax Rotations: Interdependent Self-Construal

Items	*Maximum likelihood factors*				*Principal components factors*			
	1	*2*	*3*	*4*	*1*	*2*	*3*	*4*
				With varimax				
T1	**0.595**	0.099	0.071	-0.021	**0.729**	0.109	0.065	-0.077
T2	**0.723**	0.109	0.077	0.011	**0.748**	0.176	0.139	-0.057
T3	**0.405**	**0.335**	0.140	**0.334**	**0.392**	**0.592**	-0.025	0.263
T4	**0.422**	**0.348**	-0.195	**0.322**	**0.558**	**0.395**	-0.278	**0.318**
T5	0.119	**0.392**	**0.347**	0.004	0.019	**0.404**	**0.650**	-0.094
T6	0.145	**0.695**	0.139	0.225	0.100	**0.733**	0.217	0.184
T7	0.180	**0.792**	0.131	0.244	0.162	**0.745**	0.215	0.224
T8	0.074	0.230	**0.945**	0.216	0.041	0.271	**0.778**	0.198
T9	0.018	0.150	0.137	**0.696**	0.014	0.189	0.166	**0.736**
T10	0.045	0.035	-0.044	**0.628**	-0.004	0.094	-0.112	**0.770**
T11	0.247	0.233	0.213	**0.492**	**0.338**	0.126	**0.315**	**0.588**
T12	-0.061	0.180	0.072	**0.342**	-0.226	0.250	0.071	**0.490**
T13	0.256	0.154	0.234	**0.332**	**0.419**	-0.058	**0.503**	**0.409**
T14	0.218	**0.345**	0.223	0.191	0.091	**0.652**	0.184	0.062
				With promax				
T1	**0.676**	-0.053	0.027	-0.142	**0.796**	0.003	0.014	-0.219
T2	**0.819**	-0.088	0.023	-0.124	**0.803**	0.060	0.084	-0.216
T3	**0.321**	0.206	0.052	0.217	**0.306**	**0.537**	-0.104	0.139
T4	**0.333**	0.289	**-0.310**	0.233	**0.507**	**0.316**	**-0.374**	0.231
T5	0.026	**0.418**	**0.308**	-0.187	-0.025	**0.375**	**0.662**	-0.234
T6	-0.079	**0.797**	0.021	-0.009	-0.020	**0.724**	0.170	0.049
T7	-0.071	**0.912**	-0.005	-0.024	0.038	**0.720**	0.159	0.083
T8	-0.001	0.035	**0.963**	0.064	-0.039	0.160	**0.774**	0.100
T9	-0.139	-0.028	0.074	**0.766**	-0.144	0.054	0.100	**0.771**
T10	-0.063	-0.134	-0.101	**0.741**	-0.153	-0.019	-0.185	**0.853**
T11	0.141	0.056	0.144	**0.455**	0.242	-0.061	0.245	**0.555**
T12	-0.183	0.149	0.030	**0.346**	**-0.368**	0.216	0.038	**0.532**
T13	0.203	-0.010	0.190	0.287	**0.383**	-0.266	**0.456**	**0.358**
T14	0.123	0.299	0.162	0.056	0.002	**0.663**	0.150	-0.069

NOTE: Loadings in bold are values greater than 0.30. Correlations among factors were shown by 1 and 2 (0.569); 1 and 3 (0.227); 1 and 4 (0.439); 2 and 3 (0.386); 2 and 4 (0.592); as well as 3 and 4 (0.308). Correlations among components were shown by 1 and 2 (0.350); 1 and 3 (0.180); 1 and 4 (0.377); 2 and 3 (0.218); 2 and 4 (0.383); as well as 3 and 4 (0.235).

that all correlations among factors/components of interdependent self-construal and embarrassibility are smaller for PCA than for EFA. This difference between PCA and EFA in magnitude of correlations among factors/components can also lead researchers who use PCA as a means to determine latent factors to incorrectly conclude that some components are relatively independent when in fact they are not. Furthermore, both tables show that EFA with an oblique rotation represents a clearer, simpler, and more interpretable structure of the measured variables than PCA with an orthogonal rotation. Despite the common perception that orthogonal rotations yield more lucid results, in some cases, results from oblique rotations are more clear and hence preferable.

It should be emphasized that in analyses of both measures, the EFA and PCA did not yield the same number of factors/components. In order to compare the loadings of the different statistical procedures, the EFA and PCA were constrained to generate the same number of factors/components. However, this exemplifies the differing solutions each analysis produces, especially when the eigenvalue criterion is used in comparison to the ML extraction method with fit indexes. Thus, conclusions or interpretations regarding factors/components would likely differ depending on which procedure was employed. Research goals need to be matched with procedures in order to avoid obscuring or inflating results.

THE USE OF EFA AND PCA IN CURRENT COMMUNICATION RESEARCH

With the possibility that inappropriate use of factor analytic procedures can result in incorrect conclusions, a question arises as to how factor analytic procedures are used in communication research. Two studies reported a frequent misuse of EFA in psychology from 1975 to 1984 (Ford, MacCallum, & Tait, 1986) and from 1991 to 1995 (Fabrigar et al., 1999). We wanted to determine if this was the case in communication research as well.

To explore the use of factor analytic procedures in communication, we reviewed articles published from 1990 to 2000 in three major communication journals: *Human Communication Research*, *Communication Monographs*, and *Communication Research*. Since these journals are considered to be leading journals in communication research, the assumption is that the articles published in these journals should show methodologically rigorous work. We examined all the articles in these journals from 1990 to 2000 to determine if any statistical analysis conducted in the articles addressed exploratory factor analytic questions. We then assessed whether they reported the use of PCA or common factor methods (e.g., maximum likelihood and principal factors) to address their factor analytic questions. Thus, each article was coded for the type of analysis, rotation procedure, and method for determining the number of factors/components. A summary of the findings is reported in Table 3. Out of the 193 (*Human Communication Research*), 243 (*Communication Research*), and 114 (*Communication Monographs*) empirical articles, we found 36, 53, and 30 articles, respectively, that used factor analysis procedures. The total of 119 articles constitutes 21.64% of the empirical studies published in the three journals, showing that exploratory factor analytic procedures are quite popular statistical analyses in communication.

Most notable in Table 3 is the dominance of the PCA procedure. Almost half of all the articles we reviewed used PCA. Although not reported in the table, we noticed that a greater number of authors reported using PCA for uncovering underlying dimensions or concepts rather than for data reduction purposes. Another finding in Table 3 is that only a small portion of the articles reported the exact type of EFA used (e.g., ML, principal factor methods). A large percentage (32%) of the articles did not specify the type of analysis. Since most of these articles simply reported conducting a "factor analysis" without further detail, it is impossible to determine the factor extraction method used.

Second, varimax was the dominant method used for factor/component rotation. Oblique rotation was reported in only 13 out of the 119 articles (11%) we reviewed. Additionally, some authors reported using both orthogonal and oblique rotations and chose one method over the other for simplicity or because no discernable difference was found. Approximately one fifth of the articles did not indicate their rotation methods.

Finally, when making decisions on the number of factors/components to retain, the most popular single criteria method was eigenvalues greater than one. Moreover, this method was frequently combined with the scree test (as indicated by the multiple methods category in Table 3). It was very rare for the scree test to be the only criterion for determining the number of factors/components. Similar to choice of rotation, about 45% of the articles failed to provide specifics about their factor/components number decision.

While reviewing the articles, we could not avoid questioning why certain procedures were chosen when it was obvious that other procedures were more appropriate. In other words, some researchers were making questionable decisions about factor analytic methods. For instance, a few researchers chose PCA or EFA when confirmatory factor analysis (CFA) seemed more appropriate. In other words, although the researchers stated their expectations about the factors to be extracted based on previous research or certain theoretical reasoning, they still conducted EFA or PCA. Similarly, researchers used PCA when EFA or CFA would have been more

TABLE 3
Summary Information of Current Practices in the Use of Factor Analysis

	Human Communication Research (*N* = 36)		Communication Research (*N* = 53)		Communication Monographs (*N* = 30)		*Total* (*N* = 119)	
	N	%	*N*	%	*N*	%	*N*	%
Type of analysis								
Principal components analysis	22	61.11	24	45.3	17	56.67	63	52.94
Common factors (EFA)[a]	1	2.78	9	17.0	4	13.33	14	11.76
Multiple methods[b]	2	5.56	1	1.9			3	2.52
Other					1	3.33	1	.84
Unknown[c]	11	30.56	19	35.8	8	26.67	38	31.93
Factor-component rotation								
Orthogonal	19	52.78	32	60.4	17	56.67	68	57.14
(Varimax)[d]	(15)	(41.67)	(30)	(56.6)	(16)	(53.33)	(61)	(51.26)
Oblique	6	1.67	6	11.3	1	3.33	13	10.92
(Oblimin)[e]	(2)	(5.56)	(1)	(1.9)	(1)	(3.33)	(4)	(3.36)
(Promax)[f]	(1)	(2.78)	(1)	(1.9)			(2)	(1.68)
Both orthogonal and oblique[g]	3	8.33	4	7.5	1	3.33	8	6.72
No rotation	1	2.78	2	3.8	3	10.0	6	5.04
Unknown	7	1.94	9	17.0	8	26.67	24	20.17

TABLE 3 Continued
Summary Information of Current Practices in the Use of Factor Analysis

	Human Communication Research (*N* = 36)		Communication Research (*N* = 53)		Communication Monographs (*N* = 30)		*Total* (*N* = 119)	
	N	%	*N*	%	*N*	%	*N*	%
Factor-component number procedure								
Eigenvalues > 1.0	10	27.78	15	28.3	8	26.67	33	27.73
Scree test			1	1.9			1	.84
Model fit			1	1.9			1	.84
Interpretability			1	1.9			1	.84
Multiple methods[h]	10	27.78	2	3.8	9	30.0	21	17.65
Other[i]	2	5.56	6	11.4	1	3.33	0	7.56
Unknown	14	3.39	27	50.9	12	43.33	53	44.54

NOTE: [a]Common factors (EFA) include extraction methods such as maximum likelihood and principal factors. [b]Some researchers reported both PCA and EFA. [c]"Unknown" refers to failure to specify or report procedures. [d]"Varimax" was clearly reported as the chosen orthogonal rotation. [e]"Oblimin" was clearly reported as the chosen oblique rotation. [f]"Promax" was clearly reported as the chosen oblique rotation. [g]Some authors conducted multiple PCAs and used different rotation techniques for some of the multiple PCAs. Other authors tried both orthogonal and oblique rotation techniques for the same PCA. [h]Some examples of "multiple methods" include any combination of eigenvalues > 1.0, scree test, parallel analysis, model fit, interpretability, and any procedure included in "other", etc. [i]Some examples of "other" include factor/component number decisions based on any of the following: previous research, a priori theoretical reasoning, only 1 factor solution obtained, removing factors with low reliabilities/ communalities, etc.

appropriate. For example, some researchers reported that they chose to conduct PCA in order to "assess the underlying concepts" or "confirm underlying dimensionality of the scale items." Additionally, varimax rotation was often chosen when oblique rotation seemed more appropriate. After using varimax rotation, a few researchers created composite variables by averaging the items loaded onto the same factor/components, and then reported correlations among the composite variables (i.e., the factors/components). If they suspected, expected, or even later discovered the substantial correlations among the factors/components, varimax rotation was likely not the best choice.

This analysis demonstrates the prevalence of inappropriate decisions regarding factor analytic procedures. About two thirds of the articles failed to report at least one of the three categories: factor extraction method, the number of factor/components retained, or type of rotation. It should be acknowledged that factor analytic procedures were not always the primary focus of these articles, and in addition some authors stated that they could be contacted for further information. Yet, because a significant proportion of articles did not report this information or made some problematic decisions regarding their analyses, the validity of the research findings reported in these articles is questionable.

DISCUSSION

Fabrigar et al. (1999) discussed the prevalence of inaccurate use of factor analytic methods in psychological research. Because communication researchers may be unfamiliar with this work, our purpose was to review the theoretical differences among factor analytic methods and provide suggestions regarding procedure choice. Additionally, we sought to determine the specific use (or misuse) of factor analysis within the field of communication by reviewing 119 articles within three major communication journals.

Overall, our examination of the use of EFA and PCA indicates that the questionable use of factor analysis is also evident in communication. Despite McCroskey and Young's (1979) advice, the use of factor analysis has not greatly improved over the last 22 years. Fabrigar et al. (1999) provided three reasons for the improper use of factor analysis in psychology. First, the majority of researchers receive insufficient training in factor analysis, and methodological articles on factor analysis tend to be too mathematically complex for most researchers. Second, researchers may simply follow tradition (e.g., "It has been done so this way," "Others also do it this way," etc.). Third, the poor use of factor analysis may be due to simply following defaults in some of the popular statistical software programs. We suspect that these reasons also apply in communication research.

It is our impression that applied statistics classes for social scientists provide minimal instruction on factor analysis techniques (e.g., one or two weeks). Furthermore, one of the widely popular software programs that communication scholars use, Statistical Package for Social Science (SPSS), has PCA as a default under the heading of factor analysis, which can be found in the data reduction category in its pull-down menu. However, these reasons do not make the inappropriate use of factor analysis excusable.

Although we have evidenced some examples of poor use (or lack of clear report) regarding factor analytic procedures, we can be optimistic about the future use of factor analysis in communication. First, some portion of the articles we reviewed reported clearly how their use of PCA and EFA fit their research goals as well as how and why they chose certain rotation methods. Second, information about EFA and PCA is now more widely available to researchers.[5] For example, without mathematical formulae and complicated explanations, Wegener and Fabrigar (2000) and Fabrigar et al. (1999) provide good discussion on the topic. Third, good software programs for factor analysis are available.[6] Even with the popular statistical software (e.g., SPSS), researchers have the ability to make wise choices and not simply rely on defaults or conventions. Thus, it is hoped that researchers, instructors of applied statistics courses, reviewers, and editors will take more responsibility in educating students as well as encouraging researchers to report their factor-analytic procedures more clearly.

NOTES

1. Modigliani's (1966) 26-item embarrassibility scale was used. The number of participants was 193 (160 female, 32 male, and 1 unidentified, undergraduate students; about 60% of the participants were categorized as Caucasian) with the mean age of 19.67 years old ($SD = 1.27$).

2. Interdependent self-construal was measured with 14 items of Kim and Leung's (1997) revised version of self-construal scale. The number of participants was 178 (134 female and 44 male undergraduate students; about 76.4% of the participants categorized as Caucasian) with the mean age of 19.95 years-old ($SD = 1.25$).

3. Based on previous research on the number of factors in embarrassibility, the three-factor solution appears to provide good interpretability. Although we are not experts on embarrassibility, we suspect that the first factor seems to represent the clear cause of embarrassment being the person embarrassed. The second factor appears to show embarrassment involving specific persons who can be affected by the action of the person embarrassed. The third factor seems to reflect empathic embarrassibility.

4. The analyses show unacceptable fit for 1-factor (RMSEA = 0.112, ECVI = 1.741) and 2-factor models (RMSEA = 0.102, ECVI = 1.536).

5. For example, although the first author of this paper mistakenly used PCA as if it was a factor analysis in one of her published articles, she now recognizes the distinctions between PCA and EFA.

6. Among others, we know of one program, Comprehensive Exploratory Factor Analysis (CEFA), which can be downloaded from a web site: http://quantrm2.psy.ohio-state.edu/browne/ (Browne, Cudeck, Tateneni, & Mels, 1998).

REFERENCES

Browne, M. W. (1968). A note on lower bounds for the number of common factors. *Psychometrika, 33*, 233.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24*, 445–455.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (1998). CEFA: Comprehensive Exploratory Factor Analysis [Computer software]. Available from http://quantrm2.psy.ohio-state.edu/browne/

Cliff, N., & Hamburger, C. D. (1967). The study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin, 68*, 430–445.

Dillon, W. R., & Goldstein, M. (1984). *Multivariate analysis: Methods and applications*. New York: John Wiley & Sons.

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.

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

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Humphreys, L. G. (1964). Number of cases and number of factors: An example where *N* is very large. *Educational and Psychological Measurement, 24*, 457.

Kim, M. S., & Leung, T. (1997). *A modified version of self-construal scale*. Unpublished manuscript, University of Hawai'i.

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

McCroskey, J. C., & Young, T. J. (1979). The use and abuse of factor analysis in communication research. *Human Communication Research, 5*, 375–382.

Modigliani, A. (1966). *Embarrassment and social influence*. Unpublished doctoral dissertation, University of Michigan.

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research, 14*, 403–414.

Velicer, W. F., Peacock, A. C., & Jackson, D. N. (1982). A comparison of component and factor patterns: A Monte Carlo approach. *Multivariate Behavioral Research, 17*, 371–388.

Wegener, D. T., & Fabrigar, L. R. (2000). Analysis and design for nonexperimental data: Addressing causal and noncausal hypotheses. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 412–450). Cambridge, UK: Cambridge University Press.

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

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.