Advancing Measurement Scales in Advertising, Communication, and Public Relations:
A Comprehensive Review of Scale Development Studies, 1960-2023

By

Hyoungkoo Khang[*], Ph.D
Professor
Department of Advertising and Public Relations
College of Communication and Information Sciences
The University of Alabama, USA
E-mail: khang@apr.ua.edu

Eyun-Jung Ki, Ph.D
Professor
Department of Advertising and Public Relations
College of Communication and Information Sciences
The University of Alabama, USA
E-mail: ki@apr.ua.edu

Jason Ziyuan Zhou, Ph.D
Assistant Professor
Department of Information Design and Corporate Communication,
Bentley University, USA
E-mail: jzhou@bentley.edu

&

Da-young Kang
Doctoral Candidate
College of Communication and Information Sciences
The University of Alabama, USA
E-mail: dkang9@crimson.ua.edu

[*] Corresponding author

**Advancing Measurement Scales in Advertising, Communication, and Public Relations:
A Comprehensive Review of Scale Development Studies, 1960-2023**

Abstract

To investigate scale development research in advertising, communication, and public relations, in this study, we analyzed 240 articles from 38 journals with a publication date up to 2023. We followed a structured method for scale construction, including measure definition, item-pool generation, expert content validity testing, sample administration, and scale validation. The findings indicate a post-1990s surge in scale development, with methodologies diversifying for innovative scale creation. In recent years, advanced techniques in scale development have emerged across these disciplines. However, only 33% of the studies included in this research used exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) together; this reveals a lack of comprehensive reporting on EFA processes. The present study underscores the need for enhanced rigor and reporting in future scale-development research and offers specific recommendations to accomplish this.

Keywords: scale development, factor analysis, advertising, communication, public relations

Introduction

Developing reliable, valid measures is crucial for high-quality research across disciplines, as such development enables the effective testing and advancement of theories. Although advanced techniques may be used in research, measurement issues can undermine research validity (Schmidt et al. 1985). Recognizing the importance of scale reliability and validity for academic progress, scholars have considered scale development in discussions spanning social/behavioral sciences, including psychology (e.g., Dawis 1987; Worthington and Whittaker 2006), marketing (e.g., Churchill 1979), and management (e.g., Hinkin 1995).

In advertising, communication, and public relations, scholars apply scales to explore various phenomena, such as opinions, attitudes, and behaviors While scale development remains an area in which discourse is nascent in the realm of the disciplines, a select group of scholars has pioneered discussions on the topic. McCroskey and Young (1979) highlighted issues with the application of factor analysis (FA) in communication research, whereas Park, Dailey, and Lemus (2002) examined exploratory FA (EFA) and principal component analysis (PCA) methodologies. These researchers observed reduction methods, retention criteria, and rotational techniques that were often improper or inadequately documented, underscoring the need for meticulous FA application. Morrison (2009) built on the foundational analysis of Park et al. (2002) by scrutinizing the application of EFA in scales across 51 communication journals. The outcome underscored that the methodological choices pertaining to factoring within communication research lacked the desired rigor.

Although previous studies have provided insight regarding FA in advertising, communication, and public relations research, they have often overlooked the full-scale development process and the comprehensive statistical analysis involved. They have mainly focused on EFA (e.g., Morrison 2009) or both EFA and PCA (e.g., Park, Dailey, and Lemus 2002); in contrast, they have neglected the growing importance of CFA for theory-based scale development.

With a view to addressing these gaps, our research contributes to the literature by detailing scale-development procedural steps, assessing methodologies in scale development studies, and offering recommendations for robust EFA and CFA applications. We aim to provide a comprehensive blueprint for creating refined, rigorous scales, thereby advancing scale-development discourse and practice.

To examine the status of new scale development in the advertising, communication, and public relations disciplines, this study poses the following research questions:

RQ1: Have scale-development studies taken adequate steps in developing new measures?

RQ2: Have EFA and CFA been applied correctly in scale-development studies in the field of advertising, communication, and public relations?

Method

Population

We conducted a detailed content analysis of peer-reviewed articles on new scale development up to 2023, examining 240 articles from 38 journals recognized by the National Communication Association and other major associations in advertising and public relations (Barry 1990; Henthorne, LaTour, and Loraas 1998; Ki and Ye 2017). This broad selection aimed to encapsulate comprehensive insights across closely allied fields.

Article extraction

The inclusion criteria focused on studies reporting the creation of new scales, excluding those testing existing ones. Our analysis began with a meticulous review of titles and abstracts from the 38 selected journals to ensure relevance to scale development. This foundational work underscored the importance of a rigorous selection process for capturing the essence of scale innovation in the advertising, communication, and public relations domains. A total of 240 articles were identified from this search.

Measures

In delineating the measures for scale development, our study highlighted several critical stages, which were as follows: initially defining the construct, ensuring content validity through expert review, conducting pilot tests before main data collection, and validating the scale via

statistical analyses, including EFA and CFA. From item generation to validation, each stage was scrutinized for methodological rigor. We examined EFA procedures across multiple dimensions—namely, sample characteristics, factorability assessment, extraction and rotation methods, and item retention criteria. Similarly, for CFA, our analysis encompassed the use of structural equation modeling (SEM), sample size adequacy, fit indices, and model-modification practices.

Intercoder reliability

Two well-trained doctoral students coded 10.3% of all articles ($n = 18$) to test intercoder reliability. Because all variables were dummy coded, the two coders tested intercoder reliability based on 92 variables. The Cohen's kappa values for all variables were in the range of .712 to 1; such values were considered acceptable.

Results

This study was designed to examine the status of scale development in the field of advertising, communication, and public relations. Table 1 lists the journal names and the number of articles derived from each journal. Of all the coded articles on scale-development studies, 75.8% were in the communications discipline ($n = 182$), followed by 15% in public relations ($n = 36$) and 9.2% in advertising ($n = 22$).

The first research question asked whether the scale-development studies followed adequate steps in developing a new measure. Among the 240 studies, 67.1% explicitly defined the targeted construct ($n = 161$), whereas 32.9% mentioned previous definitions only briefly or did not discuss definitions at all ($n = 79$). About 32.9% of the studies for which the aim was to develop a new measure failed to define the construct.

Regarding face and content validity, it was found that 25.4% of the studies invited experts to review the initial pool of items ($n = 61$), whereas 74.6% did not cover this process ($n = 179$). Prior to the main administration, 53.8% of the studies collected a small sample to pilot test the items ($n = 129$); in contrast, 46.3% directly implemented the main data collection ($n = 111$). Regarding the number of datasets collected to validate a scale, 237 studies implemented one round of data collection (98.8%), 143 implemented two rounds (59.6%), 40 involved three rounds (16.7%), and two involved more than four rounds of data collection (0.1%).

Regarding the process of validating the scale, 71 studies used EFA only, whereas 46 used CFA only; 119 used both EFA and CFA. As displayed in Figure 1, there has been a remarkable increase in the quantity of studies using EFA and CFA since the 1990s. More than 79% of studies using EFA and 93% of studies using CFA were published after 1990. However, the number of studies using PCA and FA did not change much from 1960 to 2023.

[Figure 1 near here]

[Table 1 near here]

Exploratory factor analysis (EFA)

Among the 240 new-scale-development articles evaluated, approximately 98.3% described conducting CFA or SEM and EFA ($n = 236$). About 29.6% used only EFA ($n = 71$), whereas 19.2% used only CFA ($n = 46$). The results indicate that most studies conducted data analysis based on more than one sample. Fifty-nine percent of the studies ($n = 143$) involved more than one round of data collection. Just over 16% of the studies ($n = 40$) involved more than three rounds of data collection, and approximately 1% ($n = 2$) involved more than four rounds of data collection.

Testing factorability. About 79.1% of the studies ($n = 190$) employed EFA to evaluate dimensionality. While the dimensionality of a dataset is related to the sizes of the correlations in the matrix, only four studies assessed the factorability of the correlation matrix before conducting EFA.

To justify performing an FA, it is necessary to determine whether the number of significant correlations existing among the observed items is sufficient. The three following indices are often used to detect the factorability of observed items: Kaiser–Meyer–Olkin (KMO), Bartlett's test of sphericity (Bartlett 1950), and individual measures of sampling adequacy (MSA).

The current study found that a large portion of the articles analyzed ($n = 64$, 26.7%) did not report any criteria that were used to evaluate the factorability of the correlation matrix. About 17.1% of studies used Bartlett's test of sphericity ($n = 41$), while 16.3% of studies reported KMO results ($n = 39$), and 4.6% assessed MSA ($n = 11$). Based on this finding, scale developers in the communication domain are encouraged to provide greater evidence of scale factorability prior to performing EFA.

Extraction methods. In the current study, it was found that most studies performing EFA reported the extraction method used ($n = 164$, 86.3%). Principal component analysis was the most commonly used method ($n = 91$, 37.9%), followed by principal axis factoring (PAF) ($n = 42$, 17.5%) and the machine learning (ML) method ($n = 23$, 9.6%). Four studies used multiple extraction methods. Examining the publication dates of the studies reporting extraction methods indicates that most studies using PCA were published prior to the majority of those using FA.

Rotation method criteria. For factor rotation, most of the articles analyzed discussed the rotation method ($n = 160$, 84.2%). Orthogonal rotation ($n = 87$) was found to be a more popular

method compared with oblique rotation ($n = 71$). Specifically, varimax rotation ($n = 80$) was identified as the dominant method among orthogonal rotations, although some studies did not specify which kind of orthogonal rotation they used ($n = 7$). In oblique rotation, 37 studies employed promax rotation, whereas 11 used direct oblimin rotation. Of the remaining studies, 23 did not explain their oblique rotation method, and seven employed more than one rotation method to achieve ideal factor loading. Although the studies used widely different rotation methods, most did not provide any specific reason for choosing the selected methods. Only 16 studies (6.9%) mentioned their selection criteria.

Factor-retention criteria. While there is no precise solution available for determining the number of factors to retain, researchers have applied several criteria, including the eigenvalue, scree plot, parallel analysis, and Velicer's minimum average partial (MAP) test.

Of the articles examined in this study, 84.7% ($n = 161$) mentioned their criteria for how many factors should be extracted. The minimum proportion of variance was, most commonly, accounted for by the factor of eigenvalue ($n = 105$), followed by retention standard ($n = 76$), scree plot ($n = 59$), parallel analysis ($n = 11$), and the MAP test ($n = 1$). Three studies used other retention standards.

Item deletion or retention criteria. When using EFA, examining item loading values and cross-loadings on the factors is the most common practice used to determine whether items should be deleted or retained. Most articles reviewed for this study explicated their deletion criteria (91.1%, $n = 173$); the majority of those studies made decisions based on loadings ($n = 134$), followed by cross-loadings ($n = 88$), commonalities ($n = 17$), and item analysis ($n = 9$). Fifteen studies used other criteria. Table 2 provides the characteristics of EFA in scale-development studies.

[Table 2 near here]

Confirmatory factor analysis (CFA)

Structural equation modeling (SEM) versus common factor analysis (FA). Among all the studies, 38.8% ($n = 93$) employed SEM-based CFA, whereas 30.4% ($n = 73$) only used FA to confirm the final scale. Three types of SEM approaches were mentioned in these articles: single model ($n = 80$), competing model ($n = 53$), and nested model ($n = 30$).

Overall fit. In reporting fit statistics, RMSEA ($n = 131$) was most frequently reported, followed by CFI ($n = 130$), Chi-square ($n = 117$), NNF/TLI ($n = 63$), RMSEA with confidence interval ($n = 48$), SRMR ($n = 58$), Chi-square/df ratio ($n = 46$), GFI ($n = 21$), NFI ($n = 25$), AIC ($n = 13$), AGFI ($n = 12$), IFI ($n = 11$), RMR ($n = 5$), BIC ($n = 3$), ECVI ($n = 1$), and RNI ($n = 1$). Forty percent of studies that employed CFA ($n = 75$) provided recommended cutoffs from previous research.

[Table 3 near here]

Validity and reliability

In the reviewed articles, various types of validity were discussed, including the following: construct validity ($n = 102$), discriminant validity ($n = 86$), convergent validity ($n = 84$), face validity ($n = 56$), content validity ($n = 55$), predictive validity ($n = 46$), and concurrent validity ($n = 37$). In terms of reliability, the overwhelming majority of studies reported internal consistency ($n = 216$), followed by test–retest reliability ($n = 9$) and alternative-form reliability ($n = 2$).

[Table 4 near here]

Discussion and Conclusions

In the current study, we scrutinized the methodologies prevalent in scale development within the realms of advertising, communication, and public relations. Anchored by the five

stages recommended by scale-development authorities, our focus was on the application of EFA and CFA techniques.

The significance of rigorous scale development for advancing academic inquiry is undeniable. We observed a notable intensification in the attention directed toward scale development since the 1990s, marked by the adoption of diverse and sophisticated methodologies. In particular, the last decade has seen a shift toward more refined scale-development techniques in advertising, communication, and public relations studies, including a move away from PCA toward SEM for confirmatory analysis. Moreover, there has been a considerable increase in the simultaneous use of EFA and CFA, indicating a maturation in the field's methodological approaches.

In addition to the advancements identified, our analysis revealed areas that were ripe for improvement. Notably, while many studies provided clear construct definitions, a significant portion either neglected this crucial step or addressed it inadequately. Precise construct definition is foundational for research in that it sets the stage for all subsequent scale-development efforts. Therefore, it is important to address this issue.

The gaps in the literature that were highlighted in this study were discrepancies in the application of EFA and CFA, unclear justifications for rotation choices, and inadequate reporting on item retention decisions. The underutilization of CFA, despite its value in refining measurement models after EFA, suggests an area in which methodological strengthening is desirable, as highlighted by Brown (2006). The field's literature often lacks transparency in EFA processes, overlooking the necessity for a methodological rationale; this critique has been echoed by Pett, Lackey, and Sullivan (2003), who also emphasized the importance of pre-EFA data factorability checks.

The choice between orthogonal and oblique rotations—often made without clear justification—also deserves attention, especially given the frequent intercorrelations among factors within communication studies. Oblique rotation, which allows for inter-factor correlations, appears most suitable for this discipline.

Another significant concern is the insufficient discussion on validity and reliability within the scale development literature. A relatively small proportion of the studies (24%) that we reviewed addressed construct or face validity, underscoring a need for a broader, more comprehensive approach to validating scales beyond quantitative assessments alone. Pett, Lackey, and Sullivan (2003) contended that validity should not be assessed based only on quantitative outcomes. It is essential that items earmarked for measuring a construct be both representative and pertinent. In this vein, content-related validity is a pivotal facet of the broader concept of validity (Cronbach, 1971).

While this study provides valuable insights into scale development processes and identifies methodological enhancements, it should be acknowledged that it has limitations. These include its focus on a select set of journals and the potential omission of emerging insights or alternative perspectives. Despite these constraints, this study's contributions involve the enhancement of how rigor and quality are measured within advertising, communication, and public relations research, and we continue to advocate for a more methodical and transparent approach to scale development.

References

Barry, Thomas E. 1990. "Publication productivity in the three leading US advertising journals: Inaugural issues through 1988." *Journal of Advertising* 19 (1): 52-60.

'Bartlett, M. S. 1950. "Tests of Significance in Factor Analysis." *British Journal of Psychology-Statistical Section* 3: 77-85. <Go to ISI>://WOS:000204954200001.

Churchill, G. A. 1979. "A paradigm for developing better measures of marketing constructs." *Journal of Marketing Research*: 64-73.

Dawis, R. V. 1987. "Scale-Construction." *Journal of Counseling Psychology* 34 (4): 481-489. https://doi.org/Doi 10.1037//0022-0167.34.4.481. <Go to ISI>://WOS:A1987K470600015.

Henthorne, Tony L, Michael S LaTour, and Tina Loraas. 1998. "Publication productivity in the three leading US advertising journals: 1989 through 1996." *Journal of Advertising* 27 (2): 53-63.

Hinkin, T. R. 1995. "A review of scale development practices in the study of organizations." *Journal of management* 21 (5): 967-988.

Ki, Eyun-Jung, and Lan Ye. 2017. "An assessment of progress in research on global public relations from 2001 to 2014." *Public Relations Review* 43 (1): 235-245.

McCroskey, James C, and Thomas J Young. 1979. "The use and abuse of factor analysis in communication research." *Human Communication Research* 5 (4): 375-382.

Morrison, John T. . 2009. "Evaluating factor analysis decisions for scale design in communication research." *Communication Methods and Measures* 3 (4): 195-215.

Park, Hee Sun, René Dailey, and Daisy Lemus. 2002. "The use of exploratory factor analysis and principal components analysis in communication research." *Human Communication Research* 28 (4): 562-577.

Pett, Marjorie A, Nancy R Lackey, and John J Sullivan. 2003. *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. sage.

Schmidt, F. L., J. E. Hunter, K. Pearlman, and H. R. Hirsh. 1985. "Forty questions about validity generalization and meta-analysis." *Personnel Psychology* 38 (4): 697-798.

Worthington, Roger L., and Tiffany A. Whittaker. 2006. "Scale development research: A content analysis and recommendations for best practices." *The Counseling Psychologist* 34 (6): 806-838.

Figure 1: Statistical analyses to validate scales by year

Number of Studies Employing Different Statistical Analyses by Year
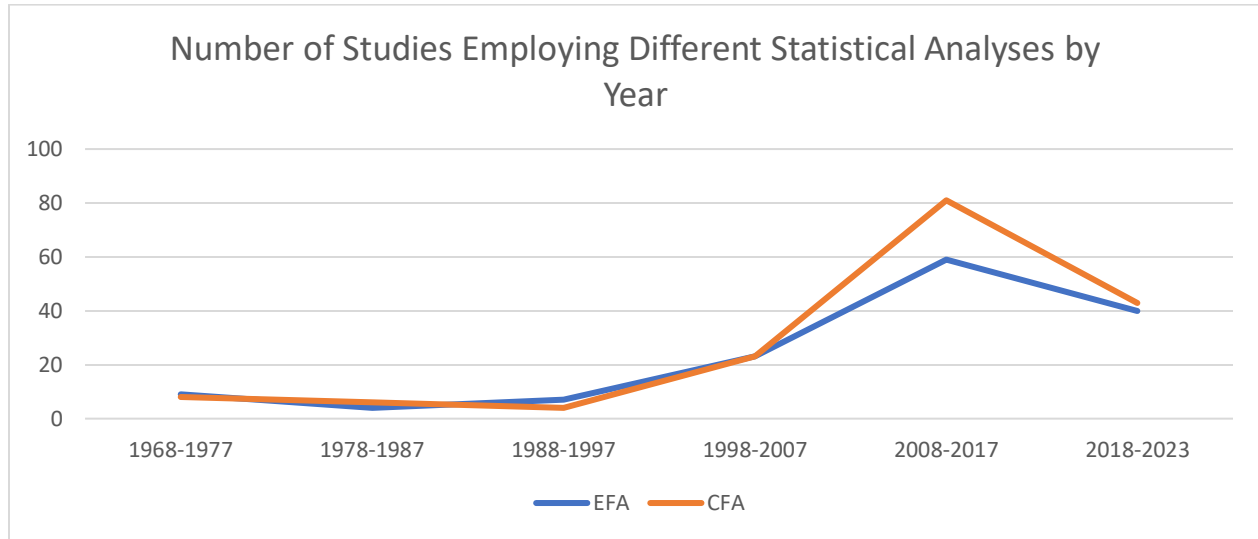
Table 1: Journals included in the analysis

Journal	Number of Articles
Journal of Advertising	4
Journal of Advertising Research	9
Journal of Current Issues and Research in Advertising	3
International Journal of Advertising	5
Communication Research	9
Human Communication Research	9
Journal of Communication	5
Journal of Broadcasting & Electronic Media	5
Journalism & Mass Communication Quarterly	16
Journal of Public Relations Research	8
Public Relations Review	14
Atlantic Journal of Communication	5
Communication Methods and Measures	14
Communication Education	18
Communication Management Quarterly	8
Communication Monographs	8
Communication Quarterly	15
Communication Reports	3
Communication Research Report	16
Communication Studies	4
Environment Communication	2
Health Communication	10
Howard Journal of Communication	1
Information, Communication & Society	4

International Journal of Media Management	1
Journal of Applied Communication Research	4
Journal of Communication Management	3
Journal of Computer-Mediated Communication	1
Journal of Family Communication	3
Journal of Health Communication	10
Journal of Intercultural Communication Research	3
Journal of International and Intercultural Communication	2
Mass Communication and Society	
Media Psychology	10
New Media & Society	4
Southern Communication Journal	2
Technical Communication Quarterly	1
Western Journal of Communication	1

Table 2: Exploratory factor analysis characteristics in scale development studies

	Number of Studies (N=240)	Percentage
Criteria Used to Assess Factorability		
KMO	39	16.3%
Bartlett's Test of Sphericity	41	17.1%
Measures of Sampling Adequacy	11	4.6%
Not Mentioned	64	26.7%
Extraction Method		
Principal Component Analysis	91	37.9%
Common Factor Analysis	70	29.2%
Principal Axis Factoring	42	17.5%
Maximum Likelihood	23	9.6%
Multiple Extraction Methods	4	1.7%
Not Reported	26	10.8%
Rotation Method		
Orthogonal Rotation	87	36.3%
Varimax	80	33.3%
Unspecified Orthogonal	7	2.9%
Oblique	71	29.6%
Promax	37	15.4%
Direct Oblimin	11	4.6%
Unspecified Oblimin	23	9.6%
Both Orthogonal and Oblique	7	2.9%
Not Reported	30	12.5%
Criteria for Factor Retention		

Eigenvalues	105	43.8%
Scree Plot	59	24.6%
Variance Accounted for by Factor	76	31.7%
MAP Test	1	0.4%
Parallel Analysis	11	4.6%
Unspecified	29	12.1%
Criteria for Item Retention		
Loadings	134	55.8%
Cross loadings	88	36.7%
Commonalities	17	7.1%
Item Analysis	9	3.8%
Others	15	7.9%
Unspecified	19	7.9%

Table 3: Confirmatory factor analysis characteristics in scale development studies

	Number of Studies (N=240)	Percentage
Confirmatory Approaches		
SEM	93	38.8%
Factor Analysis	73	30.4%
Not Applicable	74	36.6%
Typical SEM Approaches		
Single-Model Approach	80	33.3%
Competing-Model Approach	53	22.1%
Nested Model Compared	30	12.5%
Not Reported	77	32.1%
Overall Model Fit		
Chi-square	117	48.8%
Chi-square/df ratio	46	19.2%
Incremental Fit Indexes		
CFI	130	54.2%
IFI	11	4.6%
NFI	25	10.4%
NNF/TLI	63	26.3%
RNI	1	0.4%
Absolute Fit Indexes		
GFI	21	8.8%
AGFI	12	5%
RMSEA	131	54.6%
RMSEA with Confidence Intervals	48	20%
RMR	5	2.1%
SRMR	58	24.2%

Predictive Indexes		
AIC	13	5.4%
BIC	3	1.3%
Fit Indexes Criteria		
Recommended Cutoffs	75	31.3%

Table 4: Tests of reliabilities and validities

	Number of Studies (N=240)	Percentage
The Type of Validity		
Content Validity	55	22.9%
Face Validity	56	23.3%
Concurrent Validity	37	15.4%
Predictive Validity	46	19.2%
Construct Validity	102	42.5%
Convergent Validity	84	35%
Discriminant Validity	86	35.8%
The Type of Reliability		
Test-Retest Reliability	9	3.8%
Internal Consistency	216	90%
Alternative Form Reliability	2	0.8%