International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: SJIF: 5.995
Available Online at www.journalijcar.org
Volume 7; Issue 2(A); February 2018; Page No. 9631-9635
DOI: http://dx.doi.org/10.24327/ijcar.2018.9635.1600

Research Article

INVESTIGATING THE RELIABLE AND INTERPRETABLE COMPONENTS EXISTING AMONG TWENTY-ONE VARIABLES USING PRINCIPAL COMPONENT ANALYSIS

[1]Assuah, C. K., [1]Asiedu-Addo, S.K and [2]Arthur, Y.D

[1]Faculty of Science Education, Department of Mathematics Education, University of Education, Winneba
[2]College of Technology Education, Department of Interdisciplinary Studies, University of Education, Kumasi

A R T I C L E I N F O

Article History:
Received 17th November, 2017
Received in revised form 21st December, 2017
Accepted 23rd January, 2018
Published online 28th February, 2018

Key words:
Interpretable components, Uncorrelated linear combinations, geographical, Questionnaire

A B S T R A C T

The study investgated the reliable and interpretable components existing among twenty-one variables. This study used principal components analysis to obtain uncorrelated linear combinations of the original variables that account for as much of the total variance in the original variables as possible. Three hundred and fifty-four (354) elementary preservice mathematics teachers from five colleges (college 1 = 35 males and 35 females; college 2 = 35 males and 36 females; college 3= 38 males and 35 females; college 4 = 37 males and 38 females; and college 5= 33 males and 32 females) of education in Ghana participated in the study. The study identified eight (8) components or factors (exertion of authority, setting example, likeability, teacher posture, teacher attitude, student concern, teacher expectation, and setting the ground rules) as against three (3) components or factors (R = rule-based; D = dominance; N = nurturance) in the Evertson *et al.*'s (1989) study. The difference in interpretability of the components could stem from the cultural, geographical, and even economical settings that contributed for participants to respond in a particular way. This study demonstrates that even though validated questionnaire or instrument is recommended for empirical research, care should be taken when conducting research with the same questionnaire or instrument in another geographical location.

INTRODUCTION

One of the most difficult things in educational research is determining the most reliable instrument than can appropriately measure the construct(s) that a study intends to measure. Due to this difficulty, many educational researchers have often adapted or adopted instruments that have been used in similar studies to embark on their studies. One of such instruments is the questionnaire. It is a written set of questions or statements that is used to assess attitudes, opinions, beliefs, and biographical information (McMillan & Schumacher, 2006). It is the most widely used technique for obtaining information from participants. It is relatively much more economical to use, has the same questions for participants, and can ensure anonymity. It can use statements or questions, but in all cases, the participants respond to items written for specific purposes.

LITERATURE REVIEW

Factor analysis is a procedure used to determine the extent to which measurements overlap, i.e., shared variance existing among a set of variables (Williams & Brown, 1994). Its underlying purpose is to determine if measures for different variables are, in fact, measuring something in common. The procedure essentially takes the variance, as defined by intercorrelations among a set of measures, and attempts to allocate it in terms of a smaller number of underlying hypothetical variables (Williams & Brown, 1994). These underlying, hypothetical and unobservable variables are called factors. Factor analysis, therefore, is a process by which the number of variables is reduced by determining which variables "cluster" together, and factors are the groupings of variables that are measuring some common entity or construct. The main set of results obtained from a factor analysis consists of factor loadings, which are interpreted as the Pearson correlation coefficients of the original variables with factors. Like correlations, loadings range in value from -1.00 (representing a perfect negative association with the factor) through 0 to +1.00 (indicating perfect positive association). Variables typically would have loadings on all factors but would usually have high loadings on only one factor (Aron & Aron, 1999).

Communalities represent the proportion of variability for a given variable that is explained by the factors (Agresti & Finlay, 1997) and allows a researcher to examine how individual variable reflect the source of variability (Williams & Brown, 1994). Communalities may also be interpreted as the squared multiple correlation of the variable as predicted from the combination of factors, or as the sum of squared

Corresponding author:* **Assuah C. K
Faculty of Science Education, Department of Mathematics Education, University of Education, Winneba

loadings across all factors for that variable. The process by which the factors are determined from a larger set of variables is called extraction. There are actually several types of factor extraction techniques, although the most commonly used empirical approaches are principal component analysis and factor analysis (Stevens, 1992). Factor analysis is commonly used to represent the general process of variable reduction, regardless of the actual method of extraction utilized. In both principal component analysis and factor analysis, linear combinations (the factors) of original variables are produced, and a small number of these combinations typically account for the variability within the set of intercorrelations among the original variables. In principal component analysis, all sources of variability (unique, shared and error variability) are analysed for each variable. It analyses variance as opposed to covariance in factor analyses. It is the preferred method of extraction, especially when the focus of the analysis is to look for underlying structure, which is purely exploratory in nature. Its goal is to extract the maximum variance from a data set, resulting in a few orthogonal (uncorrelated) components. In this study, the following fundamental questions is addressed: (1) How many reliable and interpretable components are there among the twenty-one (21) variables in Evertson *et al.*'s (1989) conception of classroom management? (See Table 1 for all the 21 variables).

The logic behind factor analysis

The underlying mathematical objective in principal components analysis, is to obtain uncorrelated linear combinations of the original variables that account for as much of the total variance in the original variables as possible (Johnson & Wichern, 1998). These uncorrelated linear combinations are referred to as the principal components. The logic behind principal components analysis involves the partitioning of this total variance by initially finding the principal components (Stevens, 1992). The first principal component is the linear combination that accounts for the maximum amount of variance and is defined by the equation:

$$PC_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1p}x_p \qquad \text{Equation 1}$$

Where $PC_1$is the first principal component, $x_i$refers to the measure on the original variable, and a_{1i} refers to the weight assigned to a given variable (the first subscript following the a identifies the specific principal component, and the second subscript identifies the original variable). The term $a_{11}x_1$ refers to the product of the weight for variable 1 on PC_1 and the original value for an individual on variable 1. The subscript p is equal to the total number of original variables. The linear combination accounts for the maximum amount of variance within the original set of variables, the variance of the principal component is equal to the largest eigenvalue(i.e., the eigenvalue for the first component). The analytic operation then proceeds to find the second linear combination, uncorrelated with the first linear combination that accounts for the next largest amount of variance (after that which has been attributed to the first component has been resolved). The resulting equation would be:

$$PC_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2p}x_p \qquad \text{Equation 2}$$

It is important to note that the extracted principal components are not related. In other words,

$$r_{PC_1 * PC_2} = 0$$

The third component is constructed so that it is uncorrelated with the first two and accounts for the largest amount of variance in the system of original variables, after the two largest amounts have been removed. This process continues

Table 1 21 items on managing students

Rank	*Scale*	*Question*
1	R	Teachers need to be consistent with rules and consequences to get students to learn.
2	R	Talking about rules at the beginning of the year sets a positive tone for students.
3	R	Teachers should not ignore students who are inattentive but not misbehaving.
4	D	Teachers must establish their authority by laying down the law at the start of the year.
5	D	Most students will test teachers to see what they can get away with.
6	R	Setting positive expectations helps teachers limit student behaviour.
7	D	Teachers must exert their authority from the beginning.
8	N	Teachers who do not get to know their students well often have misbehavior problems.
9	R	Holding students to one set of rules still allows for their individual differences.
10	R	Teachers do not intimidate students by telling them the consequences for misbehaviour.
11	N	Treating students in a warm personal manner makes them want to behave well.
12	N	Students like teachers who let them have fun.
13	N	Students will listen to teachers they like.
14	R	Teachers should be expected to watch over students' learning and behavior all the time they are in class.
15	D	Students do not take teachers seriously if they are not stern or strict at times.
16	N	If students feel that their ideas are listened to, they do not misbehave.
17	D	Students who misbehave must be punished.
18	D	Teachers should punish the first student who misbehave as an example of the class.
19	N	Teachers who are well liked by their students do a good job of teaching.
20	N	Student behavior may indicate a lack of teacher friendliness towards students.
21	D	There is much truth to the saying "Don't smile until Christmas."

Note: R = rule-based; D = dominance; N = nurturance

METHOD

Participants

Three hundred and fifty-four (354) elementary preservice mathematics teachers from five colleges (college 1 = 70; college 2 = 71; college 3 = 73; college 4 = 75; and college 5 = 65) of education in Ghana participated in the study. College 1 consisted of 35 males and 35 females; college 2 consisted of 35 males and 36 females; college 3 consisted of 38 males and 35 females; college 4 consisted of 37 males and 38 females; and college 5 consisted of 33 males and 32 females. All the teachers were in their final stages of their education programmes and have all been introduced to effective classroom management strategies, a major component of their training.

until all variance has been accounted for by the extracted principal component.

Instrument

The instrument used in this study was a questionnaire consisting of a single task: the Managing Students Scale, a 21-item Likert-type scale, with three subscales: rule-based, dominance and nurturance conception (see Table 1 previously). The Managing Students Scale was designed to gather elementary preservice teachers' conceptions about the actions that establish and maintain control for a teacher (Evertson *et al*., 1989). The elementary preservice teachers responded on a scale ranging from strongly disagree (1) to strongly agree (5) for each of the seven items on rule-based, dominance and nurturance conception (Evertson *et al*., 1989). The questionnaire was piloted on two similar groups before it was administered to the sample of elementary preservice teachers. The reliability of the items on the questionnaire was .82, by using Cronbach's alpha. The questionnaire was administered to all the elementary preservice teachers within the last 30 minutes of a 2-hour lecture period on methods of teaching mathematics, and each participant took 15 minutes to complete.

Procedure

The first thing was to screen the data for outliers. Correlation coefficients fluctuate from sample to sample, much more so in small samples than in large. Therefore, the reliability of factor analysis is also dependent on sample size. Field (2007) suggests that over 300 participants is probably adequate, but communalities after extraction should be above 0.5. If the twenty (21) variables measure the same underlying dimension(s), then we would expect them to correlate with each other.

Variables that did not correlate with any other variables, were excluded from the data before the factor analysis was run. Conversely, variables that correlated too highly were eliminated before factor analysis was run because their existence made it impossible to determine the unique contribution to a factor of the variables that were highly correlated. Typically, principal components analysis is exploratory in nature. It determines the appropriate number of components to retain using four criteria in the decision-making process: 1) Eigenvalue-components with eidenvalues greater than 1 should be retained. This criteria is fairly reliable when the number of variables < 30 and the communalities are $> .70$, or the number of participants > 250 and the mean communality is $\geq .60$: 2) Variance-Retain components that accounted for at least 70% total variance; 3)Scree Plot-Retain all components within the sharp descent, before eigenvalues level off. This criterion is fairlyreliablewhen the number of participants is 250 and communalities $> .30$; 4) Residuals-Retain the components generated by the model if only a few residuals exceed .05

Assumptions and Limitations

The assumptions are formally stated as follows: All variables, as well as all linear combinations of variables, were normally distributed (assumption of multivariate); the relationships among all pairs of variables were linear. A major limitation for factor analysis is that correlation coefficients have a tendency to be less reliable when estimated from small samples. Therefore, if unreliable correlations exist among variables and those variables are subject to factor analysis, the resultant factors would also not be reliable.

Figure 2 Table of total variance for Eight-Component Solution

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.278	10.848	10.848	2.278	10.848	10.848
2	2.241	10.672	21.520	2.241	10.672	21.520
3	1.568	7.467	28.987	1.568	7.467	28.987
4	1.509	7.186	36.172	1.509	7.186	36.172
5	1.444	6.876	43.049	1.444	6.876	43.049
6	1.284	6.115	49.164	1.284	6.115	49.164
7	1.219	5.805	54.968	1.219	5.805	54.968
8	1.152	5.488	60.456	1.152	5.488	60.456
9	.999	4.757	65.213			
10	.948	4.513	69.725			
11	.891	4.244	73.969			
12	.804	3.831	77.800			
13	.739	3.519	81.319			
14	.684	3.258	84.577			
15	.640	3.047	87.624			
16	.586	2.790	90.414			
17	.543	2.584	92.997			
18	.479	2.279	95.276			
19	.371	1.769	97.045			
20	.333	1.587	98.632			
21	.287	1.368	100.000			

Extraction Method: Principal Component Analysis.

RESULTS

Factor analysis was conducted to determine what, if any, underlying structures existed for measures on the twenty-one variables in Table 1.

The application of the eigenvalue criterion was inappropriate because only three of the communalities were greater than 0.7, although the total number of variables were less than 30 (see Figure 2).

Figure 3 Table of Communalities

	Item	R1	R2	R3	R4	R5	R6	R7	D1	D2	D3	D4
communalities	Initial	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Extraction	.366	.615	.705	.658	.518	.628	.657	.557	.565	.735	.641
Communalities	Item	D5	D6	D7	N1	N2	N3	N4	N5	N6	N7	
	Initial	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
	Extraction	.674	.611	.621	.579	.653	.464	.586	.703	.594	.565	

Prior to analysis, three outliers were eliminated. The analysis produced an eight-component solution, which was evaluated with the following four criteria: eigenvalue, variance, scree plot, and residuals. By applying these methods of interpretation and using Varimax rotation method, the eigenvalues examined in the table of total variance (see Figure 1), indicated eight components with eigenvalues greater than 1.

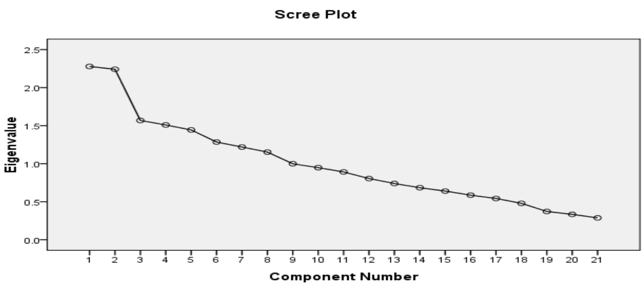

Figure 4 Scree plot

After rotation, the first, second, third, fourth, fifth, sixth, seventh, and eight components accounted for 8.53%, 8.36%, 7.79%, 7.67%, 7.45%, 7.38%, 6.86%, and 6.42% respectively of total variance of 60.45% in the original variables. The scree plot was then assessed and indicated that the eigenvalues after the eighth component somehow leveled off. Evaluation of residuals indicated that 102 (48.0%) were greater than .05. Apart from the scree plot, the other three methods of interpretation were all violated. Increasing the number of components to 10 showed that 69.79% of total variance was accounted for by the components, while the residuals reduced to 97. However, the rotated component matrix indicating the factor loadings showed that a variable loaded onto two components, which truly defeated the purpose of data reduction. Hence the eight component model was duly retained.

Table 5 Component Loadings

Components	Loading
Component 1: Exertion of Authority	
D3, Teachers must exert their authority from the beginning.	.792
D1, Teachers must establish their authority by laying down the law at the start of the school year.	.591
R7, Teachers should be expected to watch over students' learning and behavior all the time they are in class.	.589
R1, Teachers need to be consistent with rules and consequences to get students to listen.	.450
Component 2: Setting Example	
D6, Teachers should punish the first student who misbehaves as an example of the class.	.754
D5, Students who misbehave must be punished.	.747
D4, Students do not take teachers seriously if they are not stern or strict at times.	.529
Component 3: Likeability	
N4, Students will listen to teachers they like.	.663
N3, Students like teachers who let them have fun	.578
D2, Most students will test teachers to see what they can get away with.	.506
Component 4: Teacher Posture	
N6, Teachers who are well liked by their students do a good job of teaching.	.729
N7, Student misbehavior may indicate a lack of teacher friendliness toward students.	.584
Component 5: Teacher Attitude	
D2, Most students will test teachers to see what they can get away with.	.428
D7, There is much truth to the saying, "Don't smile until Christmas.	.749
R4, Setting positive expectations helps teachers limit student misbehavior.	-.712
Component 6: Student Concern	
R6, Teachers do not intimidate students by telling them the consequences for misbehavior.	.772
N5, If students feel that their ideas are listened to, they do not misbehave.	.672
R5, Holding students to one set of rules still allows for their individual differences.	-.401
Component 7: Teacher Expectation	
R7, Teachers should be expected to watch over students' learning and behavior all the time they are in class.	.412
N1, Teachers who do not get to know their students well often have misbehavior problems.	.720
N2, Treating students in a warm personal manner makes them want to behave well.	-.662
Component 8: Setting the Ground Rules	
R3, Teachers should not ignore students who are inattentive but not misbehaving.	.765
R2, Talking about rules at the beginning of year sets a positive tone for students.	.587
R5, Holding students to one set of rules still allows for their individual differences.	.470

The value of the determinant, 0.037 under the correlation matrix indicated that multicollinearity for the data was achieved since its value was greater than 0.00001. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy test indicated that with a value of 0.6091, it fell in the range of being mediocre i.e. between .5 and .7. By examining the diagonal elements of the anti-image correlation matrix, the values were all greater than .5, the bare minimum, and this was superb. The Bartlett's test of sphericityindicated that a value of 440.94 showeda significant value of .0000< 0.5.

DISCUSSION

Whereas the study by Evertson *et al*., (1989) indicated three subscales: rule-based, dominance, and nurturance conceptions that were interpretable, this study shows that eight subscales (exertion of authority, setting example, like ability, teacher posture, teacher attitude, student concern, teacher expectation, and setting the ground rules) are interpretable from the responses that were given by the 354 elementary primary preservice teachers. A difference in interpretability of the factors could stem from the cultural, geographical, and even economical settings that allowed participants to respond in a particular way. This study has clearly demonstrated that even though validated questionnaire or instrument is highly recommended for empirical research, care should be taken when conducting research with the same questionnaire or instrument in another geographical location.

References

Agresti, A. & Finlay, B. (1997). *Statistical methods for the social sciences* (3rd ed.). Upper Saddle River, N.J.: Prentice Hall.

Aron, A., & Aron, E. N. (1999). *Statistics for psychology* (2nd Ed.). New Jersey: Prentice Hall.

Everton, C. M., Emmer, E. T., Clements, B.S., Sanford, J.P., & Worsham, M. E. (1989). *Classroom management for elementary teachers*. Englewood Cliffs, NJ: Prentice-Hall.

Field, A. P. (2007). *Discovering Statistics Using SPSS* (2nd ed.). London: SAGE

Johnson, V.G. (2001). Student teachers' conceptions of classroom control. *Journal of Educational Research*, 109-117.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: evidenced-based Inquiry* (6th ed.). Boston, MA: Pearson Education, Inc.

Stevens, J.P. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ:Erlbaum.

Williams, L. J., & Brown, B. K. (1994). Method variance in organizational behavior and human resources research: Effects on correlations, path coefficients, and hypothesis testing. Organizational Behavior and Human Decision Processes, 57, 185-209.

How to cite this article:

Assuah, C. K., Asiedu-Addo, S.K and Arthur, Y.D (2018) 'Investigating the Reliable and Interpretable Components Existing among Twenty-One Variables Using Principal Component Analysis', *International Journal of Current Advanced Research,* 07(2), pp. 9631-9635. DOI: http://dx.doi.org/10.24327/ijcar.2018.9635.1600

****