INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

Exploring Factor Analysis as a Data Management Technique in Educational Survey Research

Pinky Roy[1,2], Dr. Kuki Kalpita Mahanta[1]

[1]Dibrugarh University Assam

[2]Sibsagar University, Sivasagar

DOI: https://doi.org/10.51584/IJRIAS.2026.11070187

Received: 02 August 2026; Accepted: 07 August 2026; Published: 20 August 2026

ABSTRACT

The statistical tool, factor analysis is used to extract some new factors, smaller in number, from a large amount of correlated factors to make them more manageable to further analysis. It is considered as an instrument for developing the questionnaire and proceed with that for further analysis like regression, multivariate analysis etc. Addition of more variables in the questionnaire is sometime puzzling and leads to wrong direction. So irrelevant questions can be removed and remaining to be expressed in a manageable form. The present study aims to identify the factors linked with academic performance of high school students. The Kaiser - Meyer- Olkin statistic and Bartlett's test of sphericity are applied to test the factorability of data. The number of extracted factors are determined on the basis of scree plot and total eigen value respect to each component. The orthogonal technique, varimax rotation method is applied that minimize the number of variables with high loadings on each other and try to make small loadings even smaller. The Cronbach's Alpha Coefficient measure the internal consistency of data. Convergent validity is confirmed by calculating average variance extracted > 0.4 with composite reliability > 0.6. The results confirms that application of factor analysis detect the irrelevant items and extract the valuable factors from the questionnaire which helps the researchers and administrations, teachers to take care of a few factors and make necessary actions based on that rather than considering a large amount of parameters which generally create confusions.

Key words: Factor analysis, Academic performance, Kaiser-Meyer - Olkin, Factor loading, Factor extraction

INTRODUCTION

Factor analysis is a multivariate method which reduces the complexity of the data as a whole considering the advantage of intrinsic inter dependencies. It is used to regrouping the items in the questionnaire into a limited number of clusters on the basis of shared variance. In other words, factor analysis is considered as a tool to extract some factors that explain the total variability of the group under study (Shrestha, 2021). Factor analysis represent the variability among observed related variables in-terms of lesser number of unobserved variables, so the observed variables are considered as a linear combination of unobserved factors plus error terms. (Bandalos, 2018; Yong and Pearce, 2013).

A research on education associated with different factors, which are sometime linked with each other. Only through normal view it is sometime not possible to identify the common factors and inclusion of those repeated similar items or sometime inclusion of those in non- similar group, increase the error in study. So such cases, researcher apply factor analysis to reducing the number of variables in proper manner through grouping, which provide required amount of information. The present work aims to apply factor analysis to the questionnaire item to measure the importance of factors that affect academic achievement of school level students.

The educational research always leads to the improvement of education system. The research on identification of factors affecting academic performance of students have been made from years and researchers contradict each other some time . The researchers conducted earlier shows that students learning activities are influenced by a lot of factors in school level – which includes facilities available in school, technology and teaching aid, teacher quality, textbooks, time – on – task etc. (Lubben at el., 2010] . It is also observed that school condition along with their parents socio – economic status directly effect on academic performance of students (Majumdar and Hanspal, 2024). Olubunmi, 2023 suggested that, to develop students performance and engagement, teachers and parents must work together to increase self – esteem levels, that will develop supportive school environment. So , in educational research it is seen that academic performance of students is affected by a large number of parameters. By the application of factor analysis, large number of variables are cluster into different components , where each component represent some dimension related to academic performance.

Different studies make focus on investigating the factors that affect academic achievement of students and consequences exist among them, and confirmed the existence of positive relation between socio –economic, environmental, institutional and psychological factors and performance of students in examination. Mulaudzi [2023] highlighted the significance of positive peer review relationship, good learning practices, involvement of family, accessibility of resources etc. in shaping the learning experiences of students and their academic performances. Madhukar et al., [2023] stated that students intrinsic motivation and inadequate classroom and learning facilities affect on students performance. Hassan et al. [2012] focused on factor analysis for building new factors which affect the students style of learning and assigned the factors as – attitude of students in the class, strategies imposed to comprehend the lecture, class condition, etc. Nkiruka et al., 2017, applied factor analysis and see the academic performance of primary level students and factors responsible for the variation and observed that variation mainly occur due to gender, age and environment. Yisa et al.,2023 described using factor analysis the relationship between the variables and out of 9 variables, 4 components are extracted and they accounted for 60.70% of the variation in results. The most important factors that stated under studies are motivation [Komarraju et al., 2011] , educational status of parents and students attitude(Mustaq & Khan, 2012), personality traits [Feldman et al., 2016], ethnicity [Woolf et al. 2011], age, gender(Ali et al., 2013) etc. student well being is is considered as most crucial thing to impediment academic achievement [Kutsyuruba et al., 2011]. The worth of educational institutions are linked with their students, assets of institution. The performance of students leads to the socio – economic development of a country. The good quality graduates are produced only through good performance of students. They will become the country's workforce and plays an vital role for all-round development of country [Ali et al. , 2009].

It is seen that, in recent years most of the studies on this ground are conducted among university students and less importance is given to school education. So this study, focuses on the academic performance of class X students and factors related to their performance. This is considered as the base for higher education. We use factor analysis to extract some significant factors from the items of questionnaire, that will represent maximum amount of variation in the data set. This will definitely helps the school administrators, teachers, parents, policy makers to focus on a few factors instead of concentrating on many issues.

MATERIALS AND METHODS:

2.1 Selection of sample, data collection

The proposed study was conducted in the Jorhat, Assam in India, which is considered as educational centre of Assam. We determined sample size of students by using Yamane's [1967] formula is approximately 1000 considering 3% margin of error. Out of 267 schools we consider 50 schools for the study.

The primary data were collected from 1000 class X students using questionnaire and multistage stratified random sampling is used to select the respondents for the study. Most of the responses in likert scale.

2.2 KMO and Bartlett's test:

KMO test examine the suitability of data for factor analysis. The test measure the variance proportion among all the variables. The higher proportion indicate higher KMO value results that data is more worthy to factor analysis. The calculated KMO lies between 0 and 1 and the formula is known as KMO index [Cureton & D'Aostino, 2013].

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

$$KMO = \frac{\sum_{j \neq k} {r_{jk}}^2}{\sum_{j \neq k} {r_{jk}}^2 + \sum_{j \neq k} {p_{jk}}^2}$$

Where r_{jk} is the correlation between the variable in question and p_{jk} is the partial correlation. The KMO index can be interpreted as

If the KMO value tends to zero stipulate the existence of widespread correlation, that generate problem for factor analysis. KMO value > 0.60 for any sample to be adequate for applying factor analysis. [Tabachnick and Fidell, 2013; Guttman, 1954]

Another measure named as Bartlett test is used to test that a matrix is factorable or not , tests the degree of deviation of the matrix from the identity matrix. This test is introduced by Maurice Stevenson Bartlett,1951. the correlation matrix $|R|$ to be calculated for the variables. The test statistic is defined as

$$\chi^2 = -(n - 1 - \frac{2p + 5}{6}) \times \ln|R|$$

p denotes the total variables, n is for sample size, R shows correlation matrix [Guttman, 1954]. here the p value< 0.05 indicates the applicability of factor analysis to the data set.

2.3 Procedure of Factor analysis

2.3.1 Factor analysis

Factor analysis describe the relationship among the variables, make an effort to select a new set of factors, number is lesser than original variables. Here a set of common factors are bring out, which are enough to see the relationship of original variables [Bhuyan, 2008].

A factor model can be expressed as

$$X = \Lambda F + U \qquad \text{(i)}$$

Where, $X = [X_1, X_2, X_3, \dots, X_p]', F = [f_1, f_2, f_3, \dots, f_q]', \ U = [e_1, e_2, e_3, \dots, e_p]'$

X is data vector, F represent vector of common factor, the unique factor is U.

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1q} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2q} \\ \dots & \dots & \dots & \dots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pq} \end{bmatrix}$$

Λ is a $(p \times q)$ matrix of factor loadings, where λ_{ij}'s $(i = 1,2,\dots,p; \ j = 1,2,\dots,q)$ are unknown.

And, $\text{cov}\,(U, F') = 0, E(F) = 0, V(F) = 1$, and

$$E(UU') = \begin{bmatrix} \psi_1 & 0 & 0 & \dots & 0 \\ 0 & \psi_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \psi_p \end{bmatrix}$$

We have, from (i), $\Sigma = \Lambda\Lambda' + \psi \qquad \text{(ii)}$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

If $\mathrm{cov}(F) = \phi$, Then, $\Sigma = \Lambda\phi\Lambda' + \psi$ (iii)

Σ is the variance – covariance matrix of X

From the model (i) , we get

$X_i = \sum_{j=1}^{q} \lambda_{ij} f_j + e_i , \quad i = 1, 2, \ldots, p$ (iv)

Equation (iv) represent the pattern of factor. λ_{ij} is the i^{th} variable's loading on the j^{th} factor, f_j are the factors underlying, i=1, 2, 3., ….p and j= 1, 2, 3,…,q.

2.3.2 Estimation of the factor loadings:

From model (iv), we can get,

$V(X_i) = v[\sum_{j=1}^{q} \lambda_{ij} f_j] + V(e_i)$

$\sigma_i^{\,2} = \sum_{j=1}^{q} \lambda_{ij}^{\,2} + \psi_{ij} , \quad \because V(f_j) = 1 \text{ and } \mathrm{cov}(f_j, f_j') = 0$

$\mathrm{cov}(X_i, X_i') = \mathrm{cov}[\sum_{j=1}^{q} \lambda_{ij} f_j + e_i, \sum_{j=1}^{q} \lambda_{ij}' f_j + e_i']$

$\sigma_{1i'} = \sum_{j=1}^{q} \lambda_{ij} \lambda_{ij}'$

$\therefore V(X) = \Sigma = \Lambda\Lambda' + \psi$ (vi)

$\therefore \sigma_i^{\,2} - \psi_i = \sum_{j=1}^{q} \lambda_{ij}^{\,2} = h_i^{\,2} , \; i = 1,2, \ldots, p$

$h_i^{\,2}$ is called communality.

Total communality can be defined as

$V = \sum_{i=1}^{p} h_i^{\,2}$

The factor f_j contribute to the total variance of X, as V_j

$$V_j = \sum_{i=1}^{p} \lambda_{ij}^{\,2} \Rightarrow V = \sum_{j=1}^{q} V_j$$

The proportion of variance of all the variables defined by f_j is $\frac{V_j}{V}$

This is used to obtain the number of common factors.

2.3.3 Extraction of common factor:

The extracted number of common factor is increased until explaining the " suitable proportion " of total sample variance. PCA and common factor analysis can be used for factor solutions. In the present study, principal component analysis is used to generate less number of factors which will represent the whole data set and give maximum information.

The extracted number of common factors is to be set on the basis of Kaiser's criterion (eigen value criterion) and scree plot. The eigen value of a factor determine the total variance explained by that factor. As those have

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

eigenvalues greater than 1 is retained. The reason for this is an eigenvalue greater than one is considered as significant [Pallant, 2010; Guttman, 1954; Verma, 2013].

Cattell, 1966 proposed scree plot graphs to determine the number of factors. This represents eigenvalue magnitude on vertical axis and eigenvalue number on horizontal axis. The extraction of factor should be stopped at the elbow.

2.3.4 Factor rotation:

Factors are rotated to achieve an optimal structure, that seeks to have the load of each variables on as few factors as possible and maximize the number of high loadings on that [Tabachnick, 2007]. the two approaches of factor rotation are- orthogonal and oblique factor solutions. In the present study, varimax orthogonal factor rotation method [Kaiser, 1958] is used ,

2.4 Cronbach's Alpha:

It is used to measure the reliability of the questionnaire. It follows with the assumption that more than one items measuring the same construct; in the present study different questions are asked, but combination of all that actually wants to measure students classroom Study and study behaviour. It is a measure of internal consistency. Cronbach's alpha ranges from 0 to 1 and the value > 0.70 is considered to be acceptable (Cheung et al., 2024; Lance et al., 2006)

2.5 Average variance extracted (AVE) and Composite reliability(CR):

AVE is a measurement tool that accounts the amount of variance explained by measurement error, it deals with convergent validity. The AVE and CR values ranges between 0 and 1. AVE is expressed as

$$\mathrm{AVE} = \frac{\sum_{i=1}^{n} \lambda_i^2}{n}$$

The AVE value should be at least 0.50 or more, however AVE value exceeds 0.40 is also acceptable if the CR value is adequate [Maruf et al, 2021; Fornell and Larcker, 1981].

Composite reliability indicates internal consistency in scale items, just like Cronbach's Alpha [Netemeyer, 2003]. it can be calculated as

$$\mathrm{CR} = \frac{\sum_{i=1}^{n} \lambda_i^2}{\sum_{i=1}^{n} \lambda_i^2 + \sum_{i=1}^{n} \mathrm{var}(e_i)}$$

The composite reliability value between 0.60 and 0.70 is considered as acceptable.

RESULTS AND DISCUSSIONS:

In the study, 1000 respondents are there belong to 50 different high schools of Jorhat district of Assam. Among the respondents % male and % female, 88.80% from rural area and 11.20% from urban area. Also 75% of the students studied in government schools, 21% in private school and 4% in venture school. 68.2% students take the help of tuition in their study. Among them 59.90% belong to nuclear family and 40.10% belong to joint family. Considering the father's education qualification, maximum of 33.90% HSLC passed, 22.80% HS passed, 18.90% primary school completed, 14% graduate, 6.10% have no formal education.similar type of results are seen in case of mothers. Most of the fathers 28.50% engaged in business, 27.10% doing cultivation, 10.20% teachers, 17.90% doing other jobs and 16.30% of them doing other things. According to the information, 59.30% mothers are home maker, 18.20% doing cultivation, 8.20% were teachers, 4.20% doing other job and 10% engaged in business. Now we are proceed for factor analysis.

Table 1: KMO and Bartlett's value for Student factors

KMO and Bartlett's Test		
KMO Measure		.949
Bartlett's Test of Sphericity	Approx. Chi-Square	1.364E4
	df	465
	Sig.	.000

Table (1) shows that the KMO test statistic is 0.949, which indicates that no diffusion is observed in correlation pattern and therefore the factor analysis is suitable and reliable factors should be derived from factor analysis. The p-value for Bartlett's test value < 0.000, which also indicates the strength of relationship among the variables . So wee proceed for factor analysis using this data set.

Table 2: Total Explained Variance of student factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.381	33.489	33.489	10.381	33.489	33.489	8.116	26.180	26.180
2	2.275	7.340	40.828	2.275	7.340	40.828	2.870	9.257	35.437
3	1.391	4.487	45.315	1.391	4.487	45.315	2.578	8.316	43.753
4	1.326	4.278	49.594	1.326	4.278	49.594	1.453	4.686	48.439
5	1.108	3.574	53.168	1.108	3.574	53.168	1.288	4.156	52.595
6	1.001	3.230	56.398	1.001	3.230	56.398	1.179	3.803	56.398
7	.952	3.072	59.470						
8	.893	2.882	62.352						
9	.872	2.812	65.164						
10	.833	2.688	67.852						
11	.777	2.505	70.357						
12	.727	2.344	72.700						

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

13	.709	2.286	74.987						
14	.688	2.220	77.207						
15	.648	2.089	79.296						
16	.612	1.975	81.271						
17	.588	1.898	83.169						
18	.568	1.834	85.003						
19	.518	1.669	86.672						
20	.489	1.576	88.248						
21	.464	1.498	89.746						
22	.424	1.367	91.113						
23	.412	1.330	92.443						
24	.378	1.220	93.662						
25	.344	1.109	94.772						
26	.325	1.047	95.819						
27	.318	1.025	96.844						
28	.288	.929	97.773						
29	.263	.848	98.621						
30	.252	.812	99.434						
31	.176	.566	100.000						

From table (2), we have observed that first 6 components explained a larger amount of variance, where the eigenvalues are greater than 1. After rotation, the extracted 6 factors explains in total of 56.40% of total variation. The component 1 explains 26.180% of variance ; component 2 explains 9.25% of variance; component 3,4, 5 and 6 explain 8.31%, 4.68% , 4.15% and 3.80% of variance respectively.

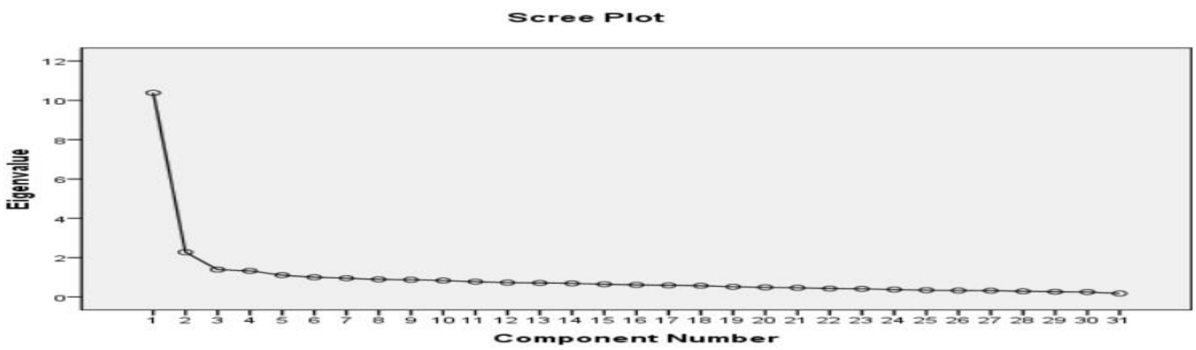

fig1: Scree plot for Student related factors

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

From fig(1), we observed that the curve seen in the scree plot become flattered from the component 5. It is seen that the eigenvalues for component 1 to 6 are more than 1 and from 7 to 31 that is less than 1. Therefore, after the extraction process, only 6 components are retained.

Table 3: Rotated component matrix for student related factors

	Component						Communalities
	1	2	3	4	5	6	Extraction
listen attentively at class(Q1), ,	.833	.157	-.075	-.104	.080	-.024	.742
Actively participate at the discussions(Q2)	.645	.288	-.114	-.068	-.139	-.080	.542
make myself prepare for examination(Q3),	.832	.154	.085	.001	-.026	-.151	.746
do my assignments regularly(Q4),	.787	.177	-.166	-.077	-.024	.005	.685
spend my vacant time in study(Q5),	.632	.100	-.022	-.076	.104	.074	.432
interested in extracurricular activities(Q6)	.356	-.037	.335	.002	-.157	-.167	.293
Have a specific place of study which I keep clean(Q7),	.739	.192	-.094	-.019	.093	-.075	.606
Study only when there is an exam(Q8)	.310	-.016	.105	-.024	.691	-.060	.589
Study only when I like(Q9),	-.163	.005	-.057	-.050	.758	.087	.615
Feel tired, bored and sleepy while study(Q10),	-.250	-.152	.607	.047	.204	.009	.497
I copy assignments from others(Q11),	-.599	-.194	.696	.053	.158	-.045	.514
I am disturbed when studying(Q12),	-.434	-.098	.576	.032	.111	-.025	.544
I am lazy(Q13),	-.401	-.004	.490	.094	.141	.207	.472
study regularly at morning and evening(Q14),	.783	.178	-.153	-.019	-.044	-.029	.671
want to get good grades(Q15),	.733	.261	-.146	-.068	.072	-.067	.641
like to study at morning(Q16),	.684	-5.753E-5	-.165	.161	.023	-.405	.686
I like to study till late night (Q17)	.450	.330	-.050	-.268	-.111	.506	.654
Prefer watching t.v(Q18)	.158	.239	.427	.155	.035	.067	.294
Feeling sleepy in class(Q19),	-.194	-.150	.636	.066	-.051	.041	.474
), Feeling hungry in class(Q20),	-.028	-.078	.660	.085	-.094	.017	.459
Difficulty in seeing(Q21),	-.070	.056	.136	.703	-.046	.082	.529
Difficulty in hearing(Q22),	-.040	-.029	.116	.760	.003	.018	.593
Come late to school(Q23),	-.080	-.154	.347	.316	-.083	.101	.267

Absent from school(Q24),	-.383	-.272	.088	.150	.125	.400	.427
have health issues (Q25).	-.168	-.188	.118	.257	.039	.649	.566
father education(Q51),	.333	.775	-.074	.015	.001	-.062	.721
Mother education(Q52),	.346	.768	-.099	-.017	-.060	-.057	.726
Parents monitoring(Q53),	.310	.656	-.083	-.058	.037	-.080	.367
Parents homework check(Q54),	.257	.434	-.130	-.019	.012	-.076	.643
electricity supply at home(Q55),	.298	.688	-.126	-.023	.044	-.058	.584
home environment(Q56).	.335	.413	-.185	-.067	-.021	-.014	.614

In the table (3) the communalities indicates the amount of variance explained by extracted components in data structure Each number presents the partial correlation coefficient between variable and rotated component. These calculated coefficients helps in reducing the number of components on which variables having high loadings > 0.40. Table (4) shows that variables Q1, Q.2, Q.3, Q.4, Q.5, Q.7, Q.14, Q.15 and Q.16 are loaded on component 1 with loadings 0.833, 0.645, 0.832, 0.787, 0.632, 0.739, -0.434, 0.783 and 0.733 respectively. Variables Q.51, Q.52 , Q.53, Q.54, Q.55 and Q.56 are loaded on component 2 with loadings 0.775, 0.768, 0.656, 0.434, 0.688 and 0.413 respectively. The variables Q.10, Q.11, Q.12, Q.13, Q.18, Q.19 and Q.20 are loaded on component 3 with loadings 0.607 , 0.576, 0.696, 0.490, 0.427, 0.636, and 0,660 respectively. Similarly, the variables Q.21 and Q.22 are loaded on component 4 with loadings 0.703 and 0.760 respectively. The variables Q.8 and Q.9 are loaded on component 5 with loadings 0.691 and 0.758. Similarly the variables Q.17, Q.24 and Q.25 are loaded with component 6 with loadings 0.506, 0.400 and 0.649 respectively. It was seen that for variables Q. 6 and Q. 23 factor loadings are less than 0.40 for all the components, so these variables are extracted from the study.

Table 4: KMO and Bartlett's value school oriented factors

KMO and Bartlett's Test		
KMO Measure		.931
Bartlett's Test of Sphericity	Approximated Chi-Square	1.088E4
	df	300
	Sig.	.000

Here the KMO value 0.931 and p-value for Bartlett's test of Sphericity < 0.000, which also indicates the strength of relationship among the variables . So factor analysis can be applied in the set.

Table 5 : Total Explained Variance school oriented factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.513	34.054	34.054	8.513	34.054	34.054	5.585	22.340	22.340
2	1.678	6.713	40.767	1.678	6.713	40.767	3.838	15.352	37.692

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

3	1.249	4.996	45.762	1.249	4.996	45.762	1.921	7.683	45.375
4	1.125	4.502	50.264	1.125	4.502	50.264	1.222	4.889	50.264
5	.977	3.908	54.173						
6	.911	3.644	57.816						
7	.838	3.353	61.169						
8	.828	3.310	64.479						
9	.756	3.023	67.503						
10	.707	2.829	70.332						
11	.676	2.705	73.036						
12	.651	2.603	75.640						
13	.647	2.590	78.229						
14	.602	2.406	80.636						
15	.591	2.362	82.998						
16	.571	2.284	85.281						
17	.546	2.184	87.465						
18	.528	2.112	89.577						
19	.498	1.994	91.571						
20	.484	1.935	93.506						
21	.444	1.776	95.281						
22	.408	1.632	96.914						
23	.391	1.565	98.479						
24	.355	1.420	99.899						
25	.025	.101	100.000						

Table (5) indicates component 1, 2, 3 and 4 accounted for 22.34% , 15.35% , 7.68% and 4.89% of variance respectively and so on. These four components are extracted with eigenvalue >1.

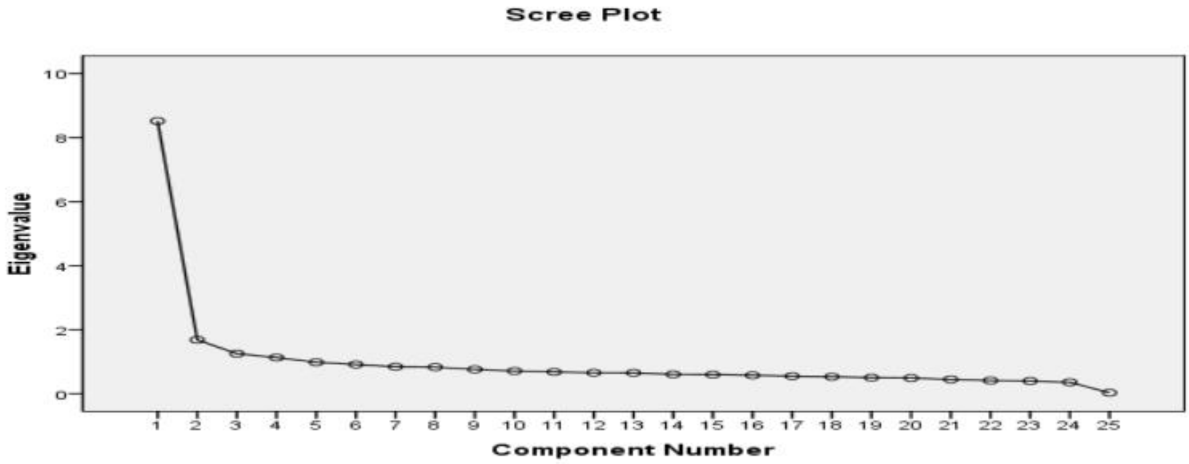

Fig2: Scree plot for school oriented factors

This scree plot presented in fig(2) is the graphical representation of the eigen values of the items in table (5). From fig(2) , the curve showed in the scree plot become flattered from the component 5.

Table 6: Rotated component matrix and communalities for school oriented factors

Rotated Component Matrix[a]					
	Component				communalities
	1	2	3	4	After extraction
teacher make good relation with students(Q26)	.614	.308	-.320	.151	.597
Smart(Q27)	.557	.232	-.163	-.120	.405
impose_discipline(Q28)	.452	.324	-.182	.069	.347
look_strict(Q29)	-.182	-.022	.700	-.164	.550
mastery_sub(Q30)	.614	.199	-.148	-.003	.438
absent_class(Q31)	-.150	-.319	.305	.305	.310
late_class(Q32)	-.264	-.381	.289	.136	.332
explain_clear(Q33)	.858	.098	-.056	-.218	.795
Updated(Q34)	.569	.336	-.055	-.131	.457
give_notes(Q35)	.659	.291	-.047	.189	.557
give_homework(Q36)	.492	.215	-.155	.144	.333
Scolds(Q37)	-.043	.008	.793	-.022	.632
Partiality(Q38)	-.130	-.061	.580	.280	.436

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VII July 2026

time schedule follow(Q39)	.506	.406	-.150	-.028	.445
overcrowd_class(Q40)	-.022	.017	-.009	.538	.604
not_waste_time(Q41)	.447	.341	-.107	-.223	.377
give equal importance to extracurricular(Q42)	.485	.480	-.009	.195	.504
conduct competitions, cultural programme frequently(Q43)	.487	.536	.004	.281	.603
library_fac(Q44)	.437	.606	-.056	.074	.566
comfortibility of class(Q45)	.385	.539	-.071	-.085	.451
behaviour of head master(Q46)	.525	.345	-.124	.020	.411
quality of academic advising of school(Q47)	.852	.091	-.056	-.221	.786
water_supply(Q48)	.247	.654	-.094	-.086	.505
Palyground(Q49)	.044	.745	.050	.022	.560
electricity_supply(Q50)	.326	.677	-.027	-.030	.566

Table(6), shows that variables Q26, Q.27, Q.28, Q.30, Q.33, Q.34, Q.35,Q.36, Q.39, Q.41, Q.42, Q.46, Q.47 are loaded on component 1 with loadings 0.614, 0.557, 0.452, 0.614, 0.858, 0.569, 0.659, 0.492, 0.506, 0.447, 0.485, 0.437, 0.525, 0.852 respectively. The variables Q.43, Q.44, Q.45, Q.48, Q.49 and Q.50 are loaded on component 2 with loadings , 0.536, 0.606, 0.539, 0.654, 0.745 and 0.677 respectively. The variables Q.29, Q.37, Q.38 are loaded on component 3 with loadings 0.0.700, 0.793, 0.580. The variables Q.40 is loaded on component 4 with loading 0.777. The variables Q.31 and Q.32 have factor loadings<0.40, so these variables are extracted from the study So we can consider the name of the components as,

Table 7: Renaming of extracted components

Student level		
components	Rename	variables
1	Classroom study and tendency towards study	listen attentively at class(Q1), Actively participate at the discussions(Q2), make myself prepare for examination(Q3), do my assignments regularly(Q4), spend my vacant time in study(Q5), Have a specific place of study which I keep clean(Q7), , study regularly at morning and evening(Q14), want to get good grades(Q15), like to study at morning(Q16),
2	Parent education and home related	father education(Q51), Mother education(Q52), Parents homework check(Q54), home environment(Q56), electricity supply at home(Q55)

www.rsisinternational.org

3	Bad habitual and problems faced in study	Feel tired, bored and sleepy while study(Q10), I copy assignments from others(Q11), I am disturbed when studying(Q12), I am lazy(Q13), Prefer watching t.v(Q18) Feeling sleepy in class(Q19), Feeling hungry in class(Q20),
4	Problems in seeing and hearing	Difficulty in seeing(Q21), Difficulty in hearing(Q22)
5	Mood to study	Study only when there is an exam(Q8), Study only when I like(Q9)
6	Personal issues	Like to study till late night(Q17), Absent from school(Q24), have health issues (Q25).
School level		
1	Teachers personality traits , teaching and school	Good relationship with students(Q26), Show smart / confident(Q27), Impose proper discipline(Q28), Mastery in subject(Q30), Explain the objects of lesion clearly(Q33), Updated with present trends(Q34), Teacher gives notes(Q35), Teachers give homework(Q36), Time schedule is followed(Q39), Class stays busy and don't waste time(Q41), School put equal importance to extracurricular activities(Q42), Behavior of the school Headmaster(Q46), Quality of academic advising of your school(Q47)
2	Logistics	School conduct competitions/ cultural programs(Q43), Library facility of school(Q44), Comfort ability of classrooms(Q45), Water supply facility(Q48), Playing ground(Q49), Electricity supply(Q50),
3	Strictness and scold	Look strict(Q29), Teacher scolds students(Q37), Teacher behaves partiality(Q38),
4	Overcrowded class	Classroom overcrowded(Q40)

Table (8) reflects that all the components have Cronbach alpha coefficient > 0.70. this indicates accuracy of instrument and reliability. From part A, it is seen that all the extracted components 1,2, 4, 5, and 6 with average extracted variance > 0.40 along with CR> 0.60 satisfied the convergent validity.

Table 8: Test for Reliability and validity :

Constructs	Cronbach's Alpha	AVE	CR
Part A: Student related			
1.Classroom Study and related factors	0.837	0.553921	0.917183
2. Parent education and home related	0.737	0.404544	0.760887
3 Bad habitual and problems faced in study	0.874	0.349525	0.786225
4, Problems in seeing and hearing	0.825	0.535905	0.697516
5. Mood to study	0.883	0.526023	0.688946

6. Personal issues	0.741	0.516479	0.761992
Part B: School related			
Teachers personality traits , teaching and school	0.767	0.518321	0.891881
logistics	0.739	0.439187	0.82331
Strictness and scold	0.875	0.54283	0.77433

In part B, all the 1st three extracted components have AVE> 0.40 and CR> 0.60. this is an indication of satisfying internal consistency and convergent validity. We exclude the 4th one because of singularity.

CONCLUSION

The present study aims to perform factor analysis of the questionnaire to identify the factors affecting academic performance of high school students. KMO and Bartlett's test of Sphericity are applied to see the applicability of factor analysis for the data set. KMO statistic 0.949 for student level data and 0.931 for school level data indicates the adequacy of sampling and factor analysis for collected data. The Bartlett's test statistic < 0.05 clearly express that correlation matrix has significant correlation among the variables. From the results carried out from the study, it was seen that factor analysis pluck the most notable factors to explain the maximum variation in the data. Application of factor analysis make it possible to convert a large amount of parameters to a manageable form. Six new components from student oriented factors and four components from school related factors are extracted applying factor analysis and varimax orthogonal factor rotation method and committed them as factors affecting academic performance of school level students, they are: student related factors:classroom study and related factors, parent education and home related factors, bad habitual and problems faced in study, problems in seeing and hearing, mood to study and personal issues; the school related factors are: teachers personality traits- teaching-school, school logistics, strictness, crowded classroom. Calculating Cronbach's alpha, AVE and CR for the extracted components give evidences for the reliability, consistency and validity of the scale items. The Application of factor analysis helps the researchers, school administration and policy creators to concentrate on fewer managable factors rather than a large amount of items and take necessary actions based on that. It will also help the researchers to apply multiple regression analysis, other multivariate approaches to managed data set.

REFERENCES

1. Agwuegbo, S. O. , Adewole, A. P. & Maduegbuna, A.N. (2010). Predicting Insurance investment:n A factor analytic approach. Journal of Mathematics and statistics, 6(3), 321-324.
2. Ahmed, S.S., Yisa, E.M., Jibrin, M., Yahaya, G. (2023). Principal component anlysis for investigating the relationship between the semester results and academic performance of students in a polytechnic in Niger State, Nigeria. Journal of applied sciences and environmental management. 27(10), 2355-2359.
3. Ali, S., Haider, Z., Munir, F., Khan, H. & Ahmed, A. (2013). Factors contributing to the students academic performance: a case study of Islamia University sub- campus. American Journal of educational research, 1(8), 283-289.
4. Bandalos, DL.(2018). Measurement theory and applications for the social sciences. The Guilford press.
5. Bartlett, M.S. (1951). The effect of standardization on chi- square approximation in factor analysis. Biometrika, 38(3/4), 337-344.
6. Cattell, R. B. (1966). The scree test for the number of factors. Multivariate behavioural research. 1, 245-276.
7. Chua, Y. P. (2009). Statistik Penyelidikan Lanjutan Ujian regresi, analysis factor dan Ujian sem. McGraw- Hill Malaysia.
8. Crosnoe, R. , Johnson, M.K., & Elder, G.H.(2004). School size and the interpersonal side of education: An examination of race/ ethnicity and organizational context. Social science quarterly, 85(5), 1259-1274.
9. Cureton, E.E., D'Agostino, R.B.(2013). Factor analysis-an applied approach. doi:10.4324/9781315799476.

10. Dhaqane, M.K., Afrah, N.A.(2016). Satisfaction of students and academic performance of benadir university. Journal of education and practice. 7(24), 59-63.
11. Feldman, G., Chandrasekhar, S.P., & Wrong, K.F.E.(2016). The freedom to excel : Belief in free will predicts better academic performance, Personality and individual differences, 90, 377-383.
12. Guttman, L., (1954). Some necessary conditions for common factor analysis. Psychometrika, 19, 149-161.
13. Hassan, S. et. al. (2012). Using factor analysis on survey study of factors affecting students learning styles, International journal of applied mathematics and informatics. 6[1]. 33-40.
14. Kaiser, H. F. . (1970). A second generation little jiffy. Psychometrika, 35, 401-415.
15. Komarraju, M., Karau, S. J. , Schmeck, R. R, Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. Personality and individual differences, 51(4), 472-477.
16. Lubben, F., Davidowitz, B., Buffler, A., Allie, S. and Scott, I.(2010). Factors influencing access student's persistence in an undergraduate science programme: A south African case study. International journal of educational development. 30, 351-358.
17. Madhukar, G. et.al.(2023). Factors affecting the stduents academic performance in secondary schools in Hyderabad. BRICS journal of educational research, 13(3), 87-89.
18. Majumder, R., Hanspal, P. (2024). Factors affecting the middle school students academic performance of Raigarh, Chhattisgarh (India) depending on socioeconomic status of their parent. Asian Journal of Education and Social Studies. 50[3], 215-218.
19. Mulaudzi, I. C. (2023), Factors affecting students academic performance: a case study of the university context, Journal of social science for policy implications.
20. Mushtaq, I., Khan, S.N., (2012). Factors affecting students academic performance. Global Journal of management and business research, 12(9).
21. Nkiruka, O. et. al.(2017). Application of factor analysis on academic performance of pupils. American journal of applied mathematics and statistics. 5[5]. 164-168.
22. Olununmi, G., Ayodele, K. (2023). Secondary school student's academic performance self esteem and school environment: an emperical assesment from Nigeria. Journal of education method and learning strategy. 1(3). 2986-9129.
23. Shrestha, N. (2021). Factor analysis as a tool for survey analysis. American Journal of applied mathematics and statistics. 9(1), 4-11.
24. Verma, J. (2013). Data analysis in management with SPSS software, Springer, India.
25. Woolf, K., Potts, H.W., & McManus, I.C. (2011). Ethnicity and academic performance in UK trained doctors and medical students: systematic review and meta- analysis. BMJ, 342, d901.
26. Yong, A.G.; Prarce, S. (2013). A biginner's guide to factor analysis: focusing on exploratory factor analysis. Tut. Quant. Meth. Psychol. 9: 79-94.