International Journal of Research and Review
[Indexed, Peer-Reviewed Journal]
Volume 13; Issue: 1; January 2026
Website: www.ijrrjournal.com
Research Paper *E-ISSN: 2349-9788; P-ISSN: 2454-2237*

Cluster Analysis of Secondary School Students' English Writing Proficiency in Bengali-Medium Schools in West Bengal

Dr. Asmita Ghosh [1], Surajit Mahato [2], Dr. Subir Sen [3]

[1] Assistant Teacher, Srinanda High School (H.S), Bolpur, Birbhum, West Bengal, India. & Former Research Scholar, Department of Education, Sidho-Kanho-Birsha University, Purulia, West Bengal, India.
[2] Research Scholar, Department of Education, Sidho-Kanho-Birsha University, Purulia, West Bengal, India.
[3] Associate Professor, Department of Education, Sidho-Kanho-Birsha University, Purulia, West Bengal, India.

Correspondence Author: Surajit Mahato

DOI: https://doi.org/10.52403/ijrr.20260114

ABSTRACT

This paper showcases a cluster analysis on the English writing proficiency level of the second-level students enrolled in the Bengali-medium schools of West Bengal, India. This paper applies a two-step clustering procedure on the SPSS software version 26. It not only helps identify the homogenous communities within the group of second-level students on the performance level of the English writing proficiency test they undertook but further probes into the linkages existing between these identified communities or level of proficiency and certain pivotal demographic characteristics: the level of schooling the children are presently in, their gender concept, geographical location (whether they belong to a rural or urban setting), and the identity of a first-generation learner. The paper required the participation of 1,204 second-level children from four districts: Birbhum, Purba Bardhaman, Paschim Bardhaman, and Murshidabad. It adopted a random stratified sampling procedure.

Keywords: Two steps Cluster, English Writing Proficiency, Bengali-Medium, first-generation learner

INTRODUCTION

Writing in a second language (L2) represents a complex cognitive activity requiring the simultaneous orchestration of multiple skills, from lexical retrieval and syntactic formulation to rhetorical organization and audience awareness. Proficiency in L2 writing is a critical academic competency, directly influencing learners' educational trajectories and future professional opportunities. However, for students in linguistically diverse contexts, such as those learning English within a Bengali-medium instructional system, achieving this proficiency presents distinct challenges. This study investigates the English writing proficiency of secondary-level students within Bengali-medium schools in the Indian state of West Bengal. By employing cluster analysis, the research aims to identify distinct, homogeneous learner profiles based on writing competency and to subsequently examine the association between these emergent groups and key socio-demographic

variables, including grade level, gender, geographical locale, and first-generation learner status. The findings are intended to provide a nuanced typology of writing proficiency, offering data-driven insights for differentiated pedagogical intervention and educational policy formulation.

Revised Literature Review:

A comprehensive understanding of L2 writing proficiency necessitates examining its constituent sub-skills, pedagogical approaches, and the contextual factors influencing its development.

English Writing Proficiency:

Dimensions and Approaches Research into L2 writing at the secondary level has explored various dimensions. A foundational study by Sharma (2009) assessed Nepalese students' proficiency in discrete grammatical features, revealing significant variability across institutions. Subsequent work has adopted more holistic frameworks. Lahuerta and Isusi Alabarte (2018), for instance, analyzed the development of complexity, accuracy, and fluency in Spanish secondary students' EFL writing, finding age-related improvements. Other studies have focused on pedagogical methodology. Ghosh et al. (2021) advocated for a process-oriented approach to address systemic weaknesses in Indian students' writing, while Vejayan and Yunus (2022), in a systematic review, highlighted persistent gaps in the adoption of varied teaching approaches (process, product, genre) in ESL classrooms. Research has also examined textual features and external influences. Xuan (2017) identified an over-reliance on specific text types in Chinese students' writing, shaped by examination culture. More recently, the impact of artificial intelligence has become salient; studies by Woo et al. (2023, 2025) have begun to explore how secondary EFL students engage with AI tools like ChatGPT, examining prompt engineering and text-editing behaviors as new facets of the writing process. In the South Asian context, Wu et al. (2013) demonstrated considerable variation in writing sub-skills among Pakistani students, correlated with school type (public/private) and location (urban/rural). This aligns with the core investigative focus of the present study on the West Bengali context. The Application of Cluster Analysis in Educational Research Cluster analysis is a valuable multivariate technique for identifying latent subgroups within a population based on shared characteristics, moving beyond average-based comparisons. Its application in educational psychology is well-established. In studies relevant to the current methodology, Sen et al. (2021) used clustering to categorize leadership styles, finding geographical location to be a strong predictor. Similarly, Saha et al. (2021) and Ansary et al. (2023) successfully employed clustering to analyze student attitudes towards yoga and value-based education, respectively, with location repeatedly emerging as a significant differentiating factor. More recently, Sen et al. (2025) demonstrated the utility of multi-model cluster analysis to profile undergraduate environmental attitudes, showing how the influence of predictors like academic stream and gender intensifies with larger cluster solutions. This methodological precedent confirms the appropriateness of cluster analysis for the present aim of classifying students into distinct writing proficiency profiles and examining the socio-demographic factors associated with each cluster.

The t-test remains a fundamental tool in hypothesis testing, with substantial recent contributions advancing its application. Key developments include work by Bauri and Mahato (2025), building on earlier foundations from Karmakar et al. (2016) and Chatterjee et al. (2016). Subsequent advances came from Mondal et al. (2018), Gayen et al. (2021), and Dandapat et al. (2021). Notable 2023 contributions include Das, Adhikari et al., Gayen and Sen, Mahato and Sen, Mondal and Saha, and Ansary. Recent work by Mahato and Das

(2024a, 2024b) continues to push methodological boundaries.

Clustering methodologies have seen significant innovation through the works of Roy and Mahato (2025), Mondal and Mahato (2025), and Sen, Mahato, Mahato, Das (2025). Das and Mahato (2024) and Das, Mahato, Sen (2023) provided important refinements, while Sen, Adhikari, Ansary and Roy (2023) and Adhikari and Sen (2023a), (2023b) established theoretical frameworks guiding current research. The works of Das and Mahato (2024), Mahato and Das (2024), and Das, Mahato, and Gayen (2024) have contributed much to the correlation methodologies and have enriched our understanding of variable relationships. The standardization techniques have evolved through works by Das et al., (2024), Mahato et al., (2024), and Das & Mahato, (2024), which have shown practical implementations across research contexts. Patra and Mahato (2025a, 2025b) have substantially developed the CFA methodologies that enhance the empirical validation of theoretical constructs. Also (Ghosh et al., 2025) conducted SEM to Validating writing skill dimensions in School Education.

The current research involves various clusters with the dichotomous variables Class, Gender, Residence, and FGL to determine the proficiency in English Writing in the Bengali Medium School. Different clusters with varying variables have been utilized for the identification of the formulation of clusters, variables, or the extent of prediction (high, mediocre, low, or very low) for the formulation of the clusters. The two-step clusters method is applied for the purpose of identifying the clusters.

The research work explores different clusters based on the Secondary School Students'English Writing Proficiency in the Bengali Medium Schools through the use of dichotomous variables. The number of different clusters is considered for the analysis of the formation of the clusters, together with their predictors and the level of prediction (high, mediocre, low, and very low) for each. The two-step clustering method is applied in the process of identifying the clusters.

Objectives of the Study:

1. To identify the cluster using research variables – Class, Gender, Residence and FGL (First Generation Learner) English Writing Proficiency in Bengali Medium School.
2. To assess predictor significance in cluster formation.

METHODOLOGY OF THE STUDY

Method: Descriptive Survey method was use by the researcher.

Population and Sample

All secondary school students enrolled in Bengali-medium schools under the West Bengal Board are included in the research. 1204 pupils from four districts – Murshidabad, Purba Bardhaman, Paschim Bardhaman, and Birbhum – make up the sample.

Sampling

Stratified random sampling used to collect the data from Respondent.

Software utilised: SPSS version 26 will be used to analyse the collected data using Two Step clustering.

RESULT AND DISCUSSION

Two step cluster

Table 1: Formation of two clusters

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	1	2
Label		
Description		
Size	60.5% (728)	39.5% (476)
Inputs	Content 3.54	Content 2.14
	Organisation 3.39	Organisation 1.86
	Vocabulary 3.21	Vocabulary 1.61
	Grammar 3.09	Grammar 1.49
	Mechanics 2.91	Mechanics 1.41
	FGL NFGL (88.6%)	FGL FGL (80.3%)
	Gender Female (65.1%)	Gender Male (60.5%)
	Reidence Urban (61.7%)	Reidence Rural (59.9%)
	Class IX (62.5%)	Class IX (56.3%)

There are two clusters. They are dominated by Female and Male students. One of them is fully comprised of rural students, while the other cluster has a majority of students from urban areas. One female-dominated cluster is fully comprised of NFGL (Non-First-Generation Learners) students, and the other cluster is dominated by more than half of FGL students. *Cluster 1 (60.5%): Mostly urban (61.7%), female (65.1%), and non-FQL (88.6%). Cluster 2 (39.5%): Mostly rural (69.9%), male (60.5%), and FQL (80.3%).*

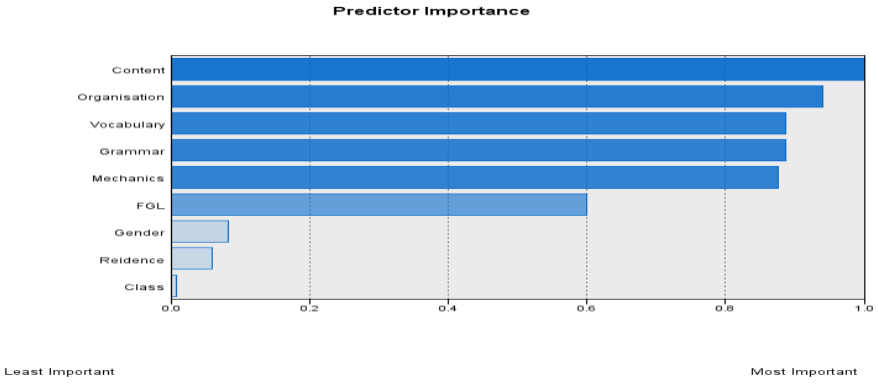

Figure 1: Predictor importance of cluster 2

Figure 1: Predictor importance for the clusters described in Table 1

From Figure 1, it is clear that Content and Organization are the most important predictors, while Grammar and Mechanics are moderately important.

Table 1. Highest Predictors: Content, Organization, Moderate Predictors: Vocabulary, Grammar, Mechanics, Lower but Notable Predictors: FGL, Gender, Residence, Least Important Predictor: Class.

Table 2: Formation of three clusters

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	3	2	1
Label			
Description			
Size	38.0% (458)	34.6% (416)	27.4% (330)
Inputs	Content 2.18	Content 3.62	Content 3.32
	Organisation 1.87	Organisation 3.46	Organisation 3.19
	Vocabulary 1.63	Vocabulary 3.29	Vocabulary 2.99
	Mechanics 1.43	Mechanics 2.97	Mechanics 2.72
	Grammar 1.52	Grammar 3.15	Grammar 2.87
	FGL FGL (84.5%)	FGL NFGL (81.2%)	FGL NFGL (100.0%)
	Gender Male (61.6%)	Gender Female (100.0%)	Gender Male (78.8%)
	Reidence Rural (60.5%)	Reidence Urban (74.5%)	Reidence Rural (54.8%)
	Class IX (55.7%)	Class IX (52.9%)	Class IX (75.2%)

There are three clusters. Two are dominated by male students and another one is female dominated. Two of them is fully comprised of rural students, while the other cluster has a majority of students from urban areas. One female-dominated cluster is fully comprised of FGL (First Generation Learners) students, and the other cluster is dominated by more than half of female non-FQL students. *Cluster 1 (27.4%): Rural (54.8%), Male (78.8%), NFGL (100%), and class IX 75.2%. Cluster 2 (34.6%): Mostly Urban (74.5%), female (100%), NFGL (81.2%) and class IX 52.9%. Cluster 3 (38.0%): Rural (60.5%), Male (61.6%), FGL and class IX (55.7%).*

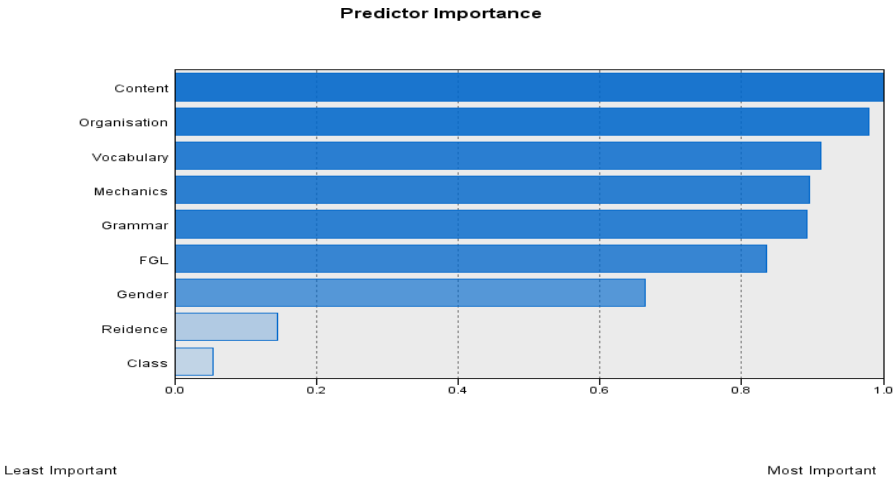

Figure 2: Predictor importance cluster 3

Figure 2: Predictor importance for the clusters described in Table 2

From Figure 2, it is clear that Content and Organization are the most important predictors, while Vocabulary, Grammar and Mechanics are moderately important.

Table 2. Highest Predictors: Content, Organization, Moderate Predictors: Vocabulary, Grammar, Mechanics, Lower but Notable Predictors: FGL, Gender, Residence, Least Important Predictor: Class.

Table 3: Formation of five clusters

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	3	4	2	5	1
Label					
Description					
Size	36.4% (438)	17.6% (212)	16.3% (196)	16.0% (193)	13.7% (165)
Inputs	Content 2.15	Content 3.47	Content 3.70	Content 3.23	Content 3.45
	Organisation 1.86	Organisation 3.27	Organisation 3.56	Organisation 3.07	Organisation 3.38
	Vocabulary 1.61	Vocabulary 3.10	Vocabulary 3.36	Vocabulary 2.89	Vocabulary 3.15
	Mechanics 1.42	Mechanics 2.75	Mechanics 3.06	Mechanics 2.60	Mechanics 2.93
	Grammar 1.51	Grammar 2.95	Grammar 3.22	Grammar 2.73	Grammar 3.08
	FGL FGL (85.8%)	FGL NFGL (76.4%)	FGL NFGL (81.1%)	FGL NFGL (99.5%)	FGL NFGL (99.4%)
	Gender Male (62.6%)	Gender Female (100.0%)	Gender Female (100.0%)	Gender Male (100.0%)	Gender Female (54.5%)
	Reidence Rural (62.6%)	Reidence Urban (100.0%)	Reidence Urban (54.6%)	Reidence Urban (81.3%)	Reidence Rural (100.0%)
	Class IX (54.8%)	Class IX (95.3%)	Class X (95.4%)	Class IX (55.4%)	Class IX (100.0%)

There are five clusters. Two are dominated by male students and another three are female dominated. Two of them are fully comprised of rural students, while the other three clusters have a majority of students from urban areas. Three female-dominated clusters are fully comprised of NFGL (Non First Generation Learners) students, and the other two male dominated clusters are dominated by more than half of male non-FQL students. Four clusters are dominated by class ix students and one cluster is dominated by class x students. *Cluster 1(13.7%): Rural (100%), female (54.5%), NFGL (99.4%), and Class ix 100%.Cluster 2 (16.3%): Mostly Urban (54.6%),female (100%), and NFGL (81.1%) and Class X 95.4% Cluster 3 (36.4%): Rural (62.6%), Male (62.6%), FGL (84.5%) and Class IX 54.8%.Cluster4(17.6%):Urban(100%),female(100%),NFGL(76.4%),Class IX(95.3%).Cluster 5 (16%) :Urban (81.3%), Male(100%), NFGL(99.5%), Class IX(55.4%)*

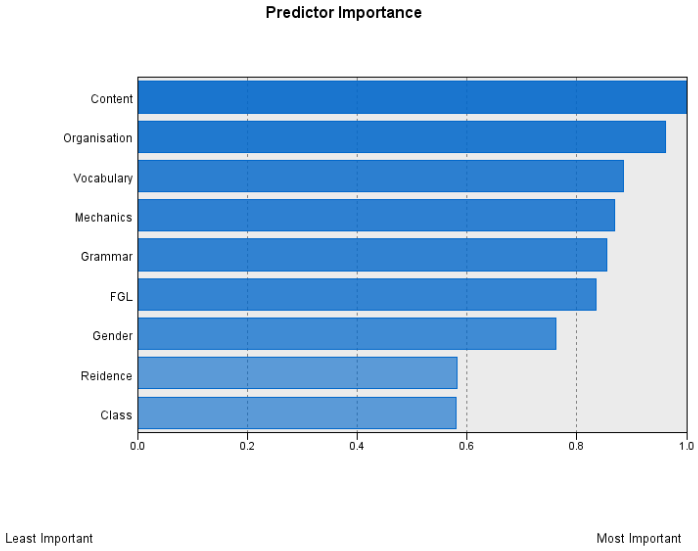

Figure 3: Predictor importance cluster 5

Figure 3: Predictor importance for the clusters described in Table 3

From Figure 3, it is clear that Content and Organization are the most important predictors, while Vocabulary, FGL, Grammar, and Mechanics are moderately important. Table 3.Highest Predictors: Content, Organization, Moderate Predictors: Vocabulary, Grammar, Mechanics, FGLLower but Notable Predictors: Gender. Least Important Predictor: Class and Residence.

Table 4: Formation of seven clusters

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	4	2	5	3	1	6	7
Label							
Description							
Size	18.2% (219)	15.8% (190)	14.9% (179)	14.5% (174)	13.7% (165)	13.3% (160)	9.7% (117)
Inputs	Gender Female (100.0%)	Gender Female (100.0%)	Gender Male (100.0%)	Gender Male (74.1%)	Gender Female (55.2%)	Gender Male (100.0%)	Gender Female (100.0%)
	Content 3.44	Content 3.68	Content 3.29	Content 2.19	Content 3.44	Content 2.04	Content 2.40
	FGL NFGL (74.4%)	FGL NFGL (85.3%)	FGL NFGL (100.0%)	FGL FGL (97.1%)	FGL NFGL (100.0%)	FGL FGL (59.4%)	FGL FGL (100.0%)
	Class IX (91.3%)	Class X (97.4%)	Class IX (58.7%)	Class IX (100.0%)	Class IX (100.0%)	Class X (86.2%)	Class X (55.6%)
	Organisation 3.22	Organisation 3.54	Organisation 3.11	Organisation 1.83	Organisation 3.37	Organisation 1.79	Organisation 2.18
	Vocabulary 3.06	Vocabulary 3.34	Vocabulary 2.96	Vocabulary 1.65	Vocabulary 3.13	Vocabulary 1.51	Vocabulary 1.88
	Mechanics 2.73	Mechanics 3.05	Mechanics 2.65	Mechanics 1.42	Mechanics 2.91	Mechanics 1.39	Mechanics 1.61
	Grammar 2.92	Grammar 3.22	Grammar 2.79	Grammar 1.55	Grammar 3.07	Grammar 1.42	Grammar 1.74
	Reidence Urban (100.0%)	Reidence Urban (56.3%)	Reidence Urban (84.4%)	Reidence Urban (58.6%)	Reidence Rural (100.0%)	Reidence Rural (65.0%)	Reidence Rural (95.7%)

There are seven clusters. Three are dominated by male students and another four are female dominated. Three are fully comprised of rural students, while the other four clusters have a majority of students from urban areas. Three female-dominated clusters are fully comprised of NFGL (Non First Generation Learners) students, and the other one female dominated cluster is dominated by more than half of female NFGL students. Four clusters are dominated by class ix students and other three clusters are dominated by class x students. *Cluster 1(13.7%): Rural (100%), female (55.2%), NFGL (100%), and Class IX 100%. Cluster 2 (15.8%): Urban (56.3%), Female (100%), NFGL (85.3%) and Class X (97.4%). Cluster 3 (14.5%): Urban (58.6%), Male (74.1%), FGL (97.1%) and Class IX (100%). Cluster 4 (18.2%): Urban(100%),Female(100%),NFGL(74.4%) and Class IX(91.3%).Cluster 5(14.9%): Urban (84.4%),Male(100%),NFGL(100%) and Class IX (58.7%).Cluster 6(13.3%): Rural(65%),Male(100%),FGL(59.4%) and Class X(65%).Cluster 7 (9.7%): Rural (95.7%),Female (100%),FGL(100%) and Class X(55.6%).*

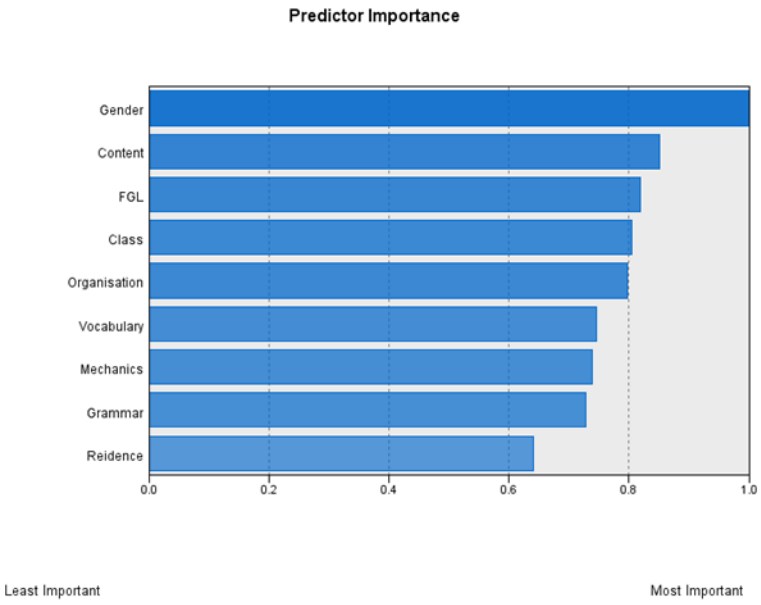

Figure 4: Predictor importance cluster 7

Figure 4: Predictor importance for the clusters described in Table 4

From Figure 4, it is clear that Gender is the most important predictors, while Content and FGLare moderately important. Table 4. Highest Predictors: Gender. Moderate Predictors: Content, FGL. Class, organisation, Vocabulary, and Mechanics are Lower but Notable Predictors: Grammar. Least Important Predictor: Residence.

Table 5: Formation of Ten clusters

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	2	1	8	4	5	6	10	9	3	7
Label										
Description										
Size	14.7% (177)	13.6% (164)	12.2% (147)	10.9% (131)	9.8% (118)	9.0% (108)	8.6% (104)	7.8% (94)	6.7% (81)	6.6% (80)
Inputs	Grammar 3.29	Grammar 3.08	Grammar 1.41	Grammar 2.21	Grammar 3.48	Grammar 2.58	Grammar 1.84	Grammar 0.57	Grammar 2.46	Grammar 3.00
	Content 3.73	Content 3.45	Content 2.03	Content 2.92	Content 3.78	Content 3.14	Content 2.54	Content 1.40	Content 2.93	Content 3.41
	Organisation 3.59	Organisation 3.38	Organisation 1.77	Organisation 2.62	Organisation 3.67	Organisation 2.97	Organisation 2.32	Organisation 0.99	Organisation 2.63	Organisation 3.24
	Mechanics 3.11	Mechanics 2.92	Mechanics 1.37	Mechanics 2.05	Mechanics 3.22	Mechanics 2.45	Mechanics 1.72	Mechanics 0.56	Mechanics 2.26	Mechanics 2.85
	Vocabulary 3.41	Vocabulary 3.15	Vocabulary 1.50	Vocabulary 2.40	Vocabulary 3.56	Vocabulary 2.78	Vocabulary 2.00	Vocabulary 0.71	Vocabulary 2.53	Vocabulary 3.08
	Gender Female (100.0%)	Gender Female (54.9%)	Gender Male (100.0%)	Gender Female (100.0%)	Gender Female (100.0%)	Gender Male (100.0%)	Gender Female (100.0%)	Gender Male (55.3%)	Gender Male (100.0%)	Gender Male (100.0%)
	Class X (97.2%)	Class IX (100.0%)	Class X (89.8%)	Class IX (72.5%)	Class IX (100.0%)	Class IX (100.0%)	Class X (57.7%)	Class IX (98.9%)	Class IX (100.0%)	Class X (100.0%)
	FGL NFGL (85.3%)	FGL NFGL (100.0%)	FGL FGL (63.9%)	FGL FGL (57.3%)	FGL NFGL (100.0%)	FGL NFGL (100.0%)	FGL FGL (100.0%)	FGL FGL (89.4%)	FGL FGL (100.0%)	FGL NFGL (98.8%)
	Reidence Urban (53.1%)	Reidence Rural (100.0%)	Reidence Rural (66.0%)	Reidence Urban (100.0%)	Reidence Urban (100.0%)	Reidence Urban (100.0%)	Reidence Rural (100.0%)	Reidence Urban (58.5%)	Reidence Rural (53.1%)	Reidence Urban (57.5%)

The cluster analysis produced ten distinct profiles. A gender-based analysis reveals a pronounced segmentation: five clusters (3, 6, 7, 8, 9) are exclusively or predominantly male-dominated, while the remaining five (1, 2, 4, 5, 10) are exclusively or predominantly female-dominated.

Geographical origin further differentiates the clusters. Three clusters (1, 3, 10) are composed entirely of rural students. Conversely, five clusters (4, 5, 6, 7, 9) are exclusively urban. The remaining two clusters (2, 8) exhibit a mixed rural-urban composition, with urban majorities in Cluster 2 (53.1%) and a rural majority in Cluster 8 (66%).

The status as a First-Generation Learner (FGL) or Non-First-Generation Learner (NFGL) also defines key groupings. Three clusters (1, 5, 6) are composed entirely of NFGL students. Another three clusters (3, 9, 10) are almost entirely composed of FGL students (89.4% to 100%). The remaining four clusters (2, 4, 7, 8) show a mixed composition, with varying majorities of either NFGL or FGL students.

Finally, class-level distribution shows that six clusters (1, 3, 4, 5, 6, 9) are dominated by Class IX students (comprising 72.5% to

100% of members). The other four clusters (2, 7, 8, 10) are dominated by Class X students (comprising 57.7% to 100% of members).

Cluster 1(13.6%): Rural (100%), female (54.9%), NFGL (100%), and Class IX 100%. Cluster 2 (14.7%): Urban (53.1%), Female (100%), NFGL (85.3%) and Class X (97.2%). Cluster 3 :(6.7%): Rural (53.1%), Male (100%), FGL (100%) and Class IX (100%). Cluster 4 (10.9%) : Urban(100%),Female(100%),FGL(57.3%) and Class IX(72.5%).Cluster 5(9.8%) : Urban (100%),Female(100%),NFGL(100%) and Class IX (100%).Cluster 6(9%): Urban(100%),Male(100%),NFGL(100%) and Class IX(100%).Cluster 7 (6.6%): Urban(57.5%),Male (100%),NFGL(98.8%) and Class X(100%).Cluster 8 (12.2%): Rural(66%),Male (100%),FGL(63.9%) and Class X(89.8%).Cluster 9(7.8%) :Urban(58.5%),Male (55.3%),FGL(89.4%) and Class IX(98.9%).Cluster 10 :(8.6%) Rural(100%),Female (100%),FGL(100%) and Class X(57.7%).

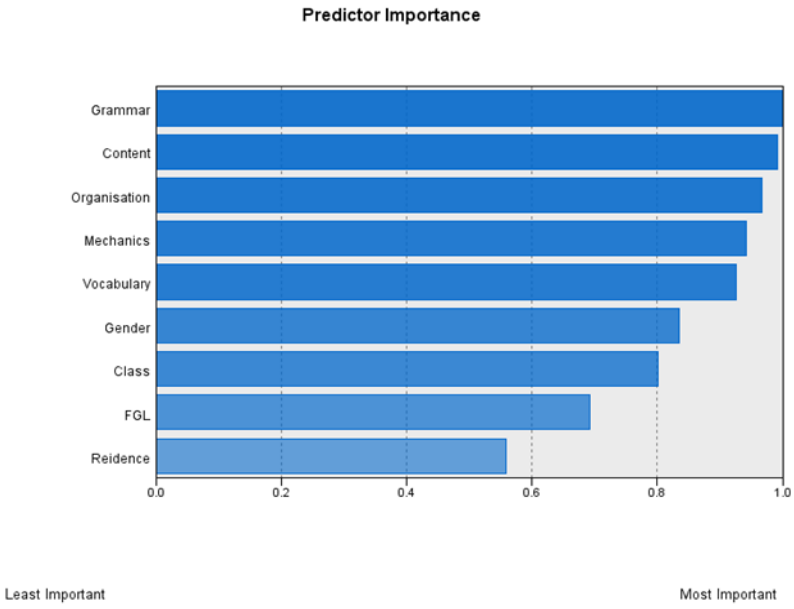

Figure 5: Predictor importance cluster 10

Figure 5: Predictor importance for the clusters described in Table 5

From Figure 5, it is clear that Grammar, Content, organisation, Vocabulary, and Mechanics are the most important predictors, while Gender and Class are moderately important.

Table 5. Highest Predictors: Grammar, Content, organisation, Vocabulary, and Mechanics. Moderate Predictors: Gender and Class. Lower but Notable Predictors: FGL. Least Important Predictor: Residence.

CONCLUSION

From Figure 1, it is clear that Content and Organization are the most important predictors, while Grammar and Mechanics are moderately important. In Table 1 Highest Predictors are Content,

Organization, and Moderate Predictors are Vocabulary, Grammar, and Mechanics. Lower but Notable Predictors are FGL, Gender, Residence, and Least Important Predictor is Class.

From Figure 2, it is clear that Content and Organization are the most important predictors, while Vocabulary, Grammar and Mechanics are moderately important. In Table 2. Highest Predictors are Content, Organization, Moderate Predictors: Vocabulary, Grammar, Mechanics, Lower but Notable Predictors are FGL, Gender, Residence, Least Important Predictor is Class.

From Figure 3, it is clear that Content and Organization are the most important predictors, while Vocabulary, FGL, Grammar, and Mechanics are moderately important. In Table 3 Highest Predictors are Content, Organization, Moderate Predictors: Vocabulary, Grammar, Mechanics, FGL. Lower but Notable Predictors is Gender. Least Important Predictors are Class and Residence.

From Figure 4, it is clear that Gender is the most important predictors, while Content and FGL are moderately important. In Table 4 .Highest Predictor is Gender. Moderate Predictors are Content, FGL. Class, Organisation. Vocabulary, Grammar. and Mechanics are Lower but Notable Predictors. Least Important Predictor is Residence.

From Figure 5, it is clear that Grammar, Content, organisation, Vocabulary, and Mechanics are the most important predictors, while Gender and Class are moderately important. In Table 5 Highest Predictors are Grammar, Content, organisation, Vocabulary, and Mechanics. Moderate Predictors are Gender and Class. Lower but Notable Predictor is FGL. Least Important Predictor is Residence.

What is known by discussing the results in terms of English writing proficiency of Bengali medium students, among the structural variables Content Organization are playing an important role in cluster formation at some point. Vocabulary Grammar and Mechanics -sometimes these are apparent important or sometimes these are apparent as moderate in terms of English writing skill. Gender is the variable which sometimes is reflected as important or sometimes it is least moderate in cluster formation. Sometimes moderation or sometimes least importantly Residence and class are used in cluster formation. All the time FGL (First Generation Learner) stands as a moderate or notable predictor in cluster formation. So, it is established that all the variables have significant importance in developing proficiency in English writing skill of Bengali medium Secondary school students of West Bengal.

Declaration by Authors
Acknowledgement: None
Source of Funding: None
Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Adhikari, A. (2023). Application of Mahalanobis Distance in Education and Educational Psychology: A Mini Review. *Innovare Journal of Education, 11*(4), 5–7.https://doi.org/10.22159/ijoe.2023v11i4.47671
2. Adhikari, A., & Sen, S. (2023a). Cluster analysis on institutional commitment and organizational climate. *International*

Journal of Research Publication and Reviews, 4(5), 4974-4988.

3. Adhikari, A., & Sen, S. (2023b). Recent trends of cluster analysis in education. *International Research Journal of Modernization in Engineering Technology and Science, 5*(8), 1858–1861.
4. Adhikari, A., Gayen, P., Mahato, R. C., Pal, I., & Sen, S. (2023e). Multi-Dimensional Data Analysis in Education: Accumulation and Comparison among Variables. *International Journal of Research Publication and Reviews,* 4(5), 2243-2245.
5. Adhikari, A., Gayen, P., Sutradhar, A., & Sen, S. (2023d). A Measure for Measure: Statistics in Education. *International Journal of Research Publication and Reviews, 4*(5), 4239-4243.
6. Adhikari, A., Gorain, S.C., Gayen, P., Pal, I., & Sen, S. (2023a). Studying the Differences: A Review on t-Test. *International Research Journal of Education and Technology, 5*(5), 338-349.
7. Adhikari, A., Mahato, R. C., & Sen, S. (2023c). Anxiety, Depression, Stress, General Self-efficacy and Specific Self-efficacy: Comparison Among Science and Social Science Students. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 3*(1), 382-385. https://doi.org/10.48175/IJARSCT-12056
8. Adhikari, A., Mahato, R.C., Gorain, S.C., & Sen, S. (2023b). A Review on Parametric and Non-Parametric Test in Education. *International Journal of Research and Analytical Reviews, 10*(2), 796-801.
9. Ahmed, E, A., Banerjee, M., Sen, S., & Chatterjee, P. (2020). Application of Mahalanobis Δ2 on Achievement Tests on Mathematics: A Study on Higher Secondary Level Students. *Indian Journal of Psychology and Education,10*(1), 36-40.
10. Ahmed, E, A., Banerjee, S., Sen. S., & Chatterjee, P. (2021).Comparison of achievement of higher secondary subjects among tribal and non-tribal students of Bodoland Territorial Region, Assam, India using Mahalanobis distance. *Journal of the Calcutta Mathematical Society, 17*(1), 61-66.
11. Ahmed, E. A., Karim, R. M., Banerjee, M., & Sen, S. (2022a). Comparison of scholastic attainment in English and Math amongst other studies at the higher secondary level: A study using Mahalanobis distance. *Kuram veUygulamadaEğitimYönetimi Educational Administration: Theory and Practice, 13*, 1-10.
12. Ahmed, E. A., Karim, R. M., Banerjee, M., Sen, S., Chatterjee, P., & Mandal, G. (2022b). A Comparative Study on Academic Achievement of Mathematics and English with Other Subjects of Secondary Level in BTR of Assam, India, Using Mahalanobis Distance. *Education Research International*.1-10.
13. Ansary, K. (2023). Study on school adjustment and academic achievement of the higher secondary school students. *The Social Science Review A Multidisciplinary Journal, 1*(2), 113-121.
14. Ansary, K., & Bauri, B. (2024). Smartphone addiction of undergraduate students with respect to gender, residence and family structure. *The Social Science Review A*

Multidisciplinary Journal, 2(5), 99-106. https://doi.org/10.70096/tssr.240205017

15. Ansary, K., Ansary, S., Adhikari, A., & Sen, S. (2023). Clustering Technique for Analyzing Attitude Towards Value-oriented Education Among Undergraduate Students. *International journal of Research Publication and reviews, 4*(5), 5576-5584.
16. Ansary, K., Saha, B., & Gorain, S. C. (2021). a study on achievement motivation of undergraduate students. *International Journal of Multidisciplinary Educational Research, 10*[9(7)], 118-121.
17. Ansary, S., Ansary, K., & Adhikari, A. (2022). Attitude Towards Social Adjustment Among the Undergraduate Students of Purulia District. *EPRA International Journal of Research and Development (IJRD), 7*(12), 21-26. http://doi.org/10.36713/epra11930
18. Bauri, L., & Mahato, S. (2025). Problematic use of mobile phone among postgraduate students in Purulia district, West Bengal. *The Social Science Review A Multidisciplinary Journal, 3*(3), 273–282. https://doi.org/10.70096/tssr.25030 3044
19. Charitaki, G., Soulis, S. G., Tzivinikou, S., &Peklari, E. (2018). Morphological skills: A key parameter in dealing with written expression and spelling in specific learning disabilities. *Creative Education, 9*(6), 879–900.
20. Chatterjee, R., Mondal, B. C., & Saha, B. (2016). Student attitudes towards using social media for educational purpose. *European Academic Research, 4*(6), 5365-5376.
21. Dandapat, M., Paramanik, S. C., Gayen, P., & Gorain, S. C. (2021). Attitude of secondary school teachers towards using ict in English classroom of Purulia district. *International Journal of Research and Analytical Reviews (IJRAR), 8*(1), 583-590.
22. Das, B. (2023). Application of Mahalanobis distance as a measure of lifestyle of health and sustainability and its components. *The Social Science Review A Multidisciplinary Journal, 1*(1), 44-52.
23. Das, B., & Mahato, S. (2024). Analysing positive mental health among students in Purulia District, West Bengal, using clustering techniques. *The Social Science Review A Multidisciplinary Journal*, *2*(3), 12-26.
24. Das, B., & Mahato, S. (2024). Lifestyle of health and sustainability: Comparison of correlations between rural-urban students in Purulia District, West Bengal using Fisher Z-transformation. *The Social Science Review A Multidisciplinary Journal*, *2*(3), 229-240.
25. Das, B., & Mahato, S. (2024). Lifestyle patterns and sustainability practices: a correlational study among undergraduate students of Purulia district, West Bengal. *International Journal of Indian Psychology, 12*(4), 1468-1488. https://doi.org/10.25215/1204.141
26. Das, B., & Mahato, S. (2024b). Lifestyle of health and sustainability: comparison of correlations between rural-urban students in Purulia district, West Bengal using fisher Z-transformation. *The Social Science Review A Multidisciplinary Journal, 2*(3), 229-240.
27. Das, B., Gayen, P., & Sen, S. (2023a). Lifestyle of health and sustainability (LOHAS): a comparative study on undergraduate students. *International Journal of Advanced Research in*

Science, Communication and Technology (IJARSCT), 3(1), 32-44.

28. Das, B., Gayen, P., Sen, S. (2023b). Lifestyle of Health and Sustainability (LOHAS) of Undergraduate Students of Purulia District of West Bengal. *EPRA International Journal of Socio-Economic and Environmental Outlook (SEEO), 10*(8), 13-19. http://doi.org/10.36713/epra0314.
29. Das, B., Mahato, S., & Gayen, P. (2024). Lifestyles of health and sustainability (LOHAS): Differentiating relationships in regard to stream of study. *The Social Science Review A Multidisciplinary Journal*, *2*(1), 1-13.
30. Das, B., Mahato, S., & Sen, S. (2023). Clustering technique for analyzing lifestyle of health and sustainability of undergraduate students. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, *3*(1), 207-221. https://doi.org/10.48175/568
31. Eckes, T. (2008). Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing, 25*(2), 155–185.
32. Gayen, P. (2024a). Cluster analysis on smartphone addiction of PG students. *The Social Science Review A Multidisciplinary Journal, 2*(5), 118-128.https://doi.org/10.70096/tssr.240205020
33. Gayen, P., & Sen, S. (2023). A study on interest in teaching of pre-service language trainee teachers of Bankura and Purulia district of West Bengal. *International Journal of Multidisciplinary Educational Research, 12*[7(1)], 79-83.
34. Gayen, P., Dandapat, M., Das, C., & Ansary, K. (2021). Attitude towards English as a language and medium of instruction: a study in secondary school students in Cooch Behar district of west Bengal. *International Journal of Research and Analytical Reviews (IJRAR),8*(1), 120-124.
35. Gayen, P., Sen, S., & Adhikari, A. (2023). Relationship between Organizational Climate and Institutional Commitment of Secondary School Teachers of West Bengal. *International Journal of Scientific Research and Engineering Development, 6*(3), 426-436.
36. Ghosh, A., Gayen, P., & Sen, S. (2021). Developing writing skill in English of secondary school students: A process approach. *International Journal of Multidisciplinary Research and Development, 8*(4), 26-29.
37. Ghosh, A., Mahato, S., & Sen, S. (2025). Validating writing skill dimensions through structural equation modelling in secondary education. *Journal of Global Research in Education and Social Science*, *19*(3), 336-347. https://doi.org/10.56557/jogress/2025/v19i39590
38. Gorain, S. C., Saha, B., Maji, S., & Sen, S. (2022). A study on relationship and cluster analysis among internet dependency, social isolation and personality. *International Journal of Research Publication and Reviews, 3*(1), 884-888.
39. Hamp-Lyons, L. (1991). *Assessing second language writing in academic contexts*. Ablex Publishing Corporation.
40. In'nami, Y., & Koizumi, R. (2016). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly, 13*(3), 250–276.

41. Karmakar, T., Paul, A., Mondal, A., & Saha, B. (2016). Intelligence in relation to height and weight among secondary school students. *American Journal of Educational Research, 4*(16), 1145-1148
42. Kaur, G., Saha, S., Das, A., & Naithani, M. (2020). Prevalence and Health related effects of internet addiction on medical students in a tertiary care centre. *Indian J Comm Health, 32*(1), 145-150.
43. Khan, S., Roy, S., Gorain, S. C., & Adhikari, A. (2023). Cyber schooling: a study on the higher education learners. *IAR Journal of Humanities and Social Sciences, 4*(1), 1-10.
44. Khatun, S., Ansary, K., & Adhikari, A. (2022). Attitude Towards Yoga Education Among Undergraduate Students. *EPRA International Journal of Multidisciplinary Research (IJMR), 8*(12), 9-13. https://doi.org/10.36713/epra11931
45. Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing, 26*(2), 275–304.
46. Kotian, S.S., & Akshara, N. (2024). Assessment of examination anxiety among secondary school students in selected private high schools in udupi. *Manipal Journal of Nursing and health Sciences, 10*(2), 21-25.
47. Lahuerta, A. C., & IsusiAlabarte, A. (2018). Development of EFL students' writing in secondary education. *Journal of Linguistics and Education Research, 1*(1), 1-13. https://doi.org/10.30564/jler.v1i1.299
48. Lee, Y. W., & Kantor, R. (2007). Evaluating prototype tasks and alternative rating schemes for a new ESL writing test through G-theory. *International Journal of Testing, 7*(4), 353–385.
49. Mahanti, J., Mondal, B. C., Saha, B. (2016). Internet Dependency of Undergraduate Students: An Empirical Study. *American International Journal of Research in Humanities, Arts and Social Sciences, 15*(2), 171-174.
50. Mahato, D., Gorain, S.C., Roy, S., & Adhikari, A. (2022). Introspecting Flipped Classroom: A Survey on Higher Education Students. *Galore International Journal of Applied Sciences and Humanities, 6*(4), 56-69. https://doi.org/10.52403/gijash.20221009
51. Mahato, R. C., & Sen, S. (2021). Application Of Mahalanobis Distance to Determine the Dynamical Nature of Academic Stress, Self Efficacy in Mathematics and Anxiety in Mathematics. *International Journal of Advances in Engineering and Management (IJAEM), 3*(5), 1398-1401.
52. Mahato, R. C., & Sen, S. (2021a). Academic Stress, Self-Efficacy and Anxiety: A Study on Mathematics of Higher Secondary Level Students in Purulia District of West Bengal, India. *International Journal of Creative Research Thoughts (IJCRT), 9*(5), c969-c980.
53. Mahato, R. C., & Sen, S. (2023). A Study of Contexts Knowledge (CK1), Technological Pedagogical Content Knowledge (TPCK) and Attitudes Towards Creative Teaching (ACT) Among the Pre-Service Mathematics Trainee Teachers in West Bengal, India. *Journal of Emerging Technologies and Innovative Research (JETIR), 10*(4), h35-h43.
54. Mahato, R.C., Sen, S., & Adhikari, A. (2023). A Study of DASS-21 and The

Self-Efficacy Scale on Post-Graduate Students. *International Journal of Research Publication and Reviews, 4*(6), 4249-4255.

55. Mahato, S., & Das, B. (2024). Comparison of environmental attitude by applying t-test and Mahalanobis distance (MD) of undergraduate students in Purulia. *The Social Science Review A Multidisciplinary Journal, 2*(5), 133-140. https://doi.org/10.70096/tssr.240205022
56. Mahato, S., & Das, B. (2024). Mental well-being among students with respect to gender, institution and residence: Insights from Purulia District, West Bengal. *The Social Science Review A Multidisciplinary Journal*, *2*(2), 164-175.
57. Mahato, S., & Das, B. (2024, August). Understand gender-specific comparison of correlations between the lifestyles of health and sustainability and its components using the Fisher Z-transformation. *International Journal of Research Publication and Reviews, 5*(8), 650–659. https://doi.org/10.55248/gengpi.5.1224.3567
58. Mahato, S., Das, B., & Gayen, P. (2024). Achievement on language subjects of secondary school students: differentiating relationships in regard to gender and type of institute. *The Social Science Review A Multidisciplinary Journal, 2*(1), 78-86.
59. Mahato, S., Das, B., & Sen, S. (2023). Test of changing status in achievement on language subject for Class VII student: A study by Mahalanobis distance. *International Journal of Research Publication and Reviews*, *4*(10), 1540-1545.
60. Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341.
61. Mohanta, R., & Mahato, S. (2025). Pioneering voices: The enduring contributions and contemporary relevance of transformative women in history. *The Social Science Review A Multidisciplinary Journal*, *3*(4), 91-102. https://doi.org/10.70096/tssr.250304016
62. Mohanta, R., Adhikari, A., Pal, I., Sen, S. (2023). Introspecting Institutional Commitment Using Cluster Analysis. *International Research Journal of Education and Technology, 5*(4), 198-217.
63. Mohanta, R., Sen, S., Adhikari, A., & Pal, I. (2023a). Perceptional Environment: A Study on Organizational Climate Using Cluster Analysis. *International Journal of Research Publication and Reviews, 4*(4), 1336-1346.
64. Mondal, N., & Saha, B. (2013). Achievement difference in science at secondary level in Darjeeling District: A comparative study. *International Journal of Scientific Research, 2*(2), 85–86. https://doi.org/10.15373/22778179/FEB2013/31
65. Mondal, S., & Mahato, S. (2025). Clustering of postgraduate student's cognitive test anxiety (CTA). *The Social Science Review A Multidisciplinary Journal, 3*(3), 72-85.https://doi.org/10.70096/tssr.250303014

66. Mukand, A. (2023). A study on internet addiction, Social Media addiction & Academic performance among college students. *International Journal of Indian Psychology, 11*(4), 755-761.
67. Patra, B., & Mahato, S. (2025a). A study on the test anxiety among higher secondary school students in Purulia. *International Journal of Sciences and Innovation Engineering, 2*(12), 129-139. https://doi.org/10.70849/ijsci12013 14801039
68. Patra, B., & Mahato, S. (2025b). Mobile phone problematic use among postgraduate students in Purulia: a confirmatory factor analysis study. *The Social Science Review A Multidisciplinary Journal, 3*(3), 294-303. https://doi.org/10.70096/tssr.250303047.
69. Roy, C., & Mahato, S. (2025). Mobile phone problematic use among postgraduate students in Purulia district of west Bengal: a descriptive study. *The Social Science Review A Multidisciplinary Journal, 3*(3), 249-260. https://doi.org/10.70096/tssr.250303040
70. Saha, B., Sen, S., & Adhikari, A. (2021). Analysis of Attitude Towards Yoga Among College Students Using Clustering Techniques. *EPRA International Journal of Multidisciplinary Research (IJMR), 7*(9), 308-314. https://doi.org/10.36713/epra8552
71. Sen, B., Mondal, N., & Saha, B. (2013). A Comparative Study of Poor Achievement in Physics at the Higher Secondary level with respect to Secondary level in Birbhum District. *International Journal of Scientific Research, 2*(4), 66-67.
72. Sen, S., & Samanta, T. K. (2015). Content knowledge and pedagogical content knowledge in the eighth grade mathematics textbook of West Bengal Board of Secondary Education. *Scholarly Research Journal for Interdisciplinary Studies, 3*(19), 619–629.
73. Sen, S., Adhikari, A., Ansary, K., & Roy, S. (2023a). Clustering technique for analysing leadership style of the head of the institutions. *International Journal of Advanced Research in Science, Communication and Technology, 3*(3), 220–228.
74. Sen, S., Gayen, P., Mahato, R. C., & Adhikari, A. (2023b). A Correlational Study on Organisational Climate and Institutional Commitment of Secondary School Teachers. *International Journal of Multidisciplinary Research and Publications (IJMRAP), 5*(12), 152-155.
75. Sen, S., Mahato, S., Mahato, R. C., & Das, B. (2025). Clustering technique for analysing environmental attitude among undergraduate students in Purulia district, West Bengal. *Journal of Social Studies,21*(1),1–12. https://doi.org/10.21831/jss.v20i1.70279
76. Sen, S., Mahato, S., Mandal, G., & Saha, B. (2025). A study of positive mental health on Bengali speaking students of higher education in India and Bangladesh. *American Journal of Educational Research*, *13*(5), 268-276. https://doi.org/10.12691/education-13-5-1
77. Sen, S., Pal, I., & Adhikari, A. (2023). Comparison among Self-Efficacy, Depression, Anxiety and Stress of Postgraduate Students by Mahalanobis Distance. *International Journal of Advanced Education and Research, 8*(1), 85-88.

78. Sharma, C. (2009). Writing proficiency of the students of higher secondary level [Unpublished master's thesis]. *Department of English Education, Tribhuvan University*. https://hdl.handle.net/20.500.14540/6417
79. Singh, A. B., Mishra, A. K., Vijay, M., Karanth, V., & Mittal, K. (2023). Prevalence and correlates of social networking addiction and its association with adult ADHD and stress among students of a medical college. *Indian Journal of Community and Family Medicine, 9*(1), 65-71.
80. Struthers, L., Lapadat, J. C., & MacMillan, P. D. (2013). Assessing cohesion in children's writing: Development of a checklist. *Assessing Writing, 18*(3), 187–201.
81. Vejayan, L., & Yunus, M. Md. (2022). Writing skills and writing approaches in ESL classroom: A systematic review. *International Journal of Academic Research in Business and Social Sciences, 12*(6), 1-16. https://doi.org/10.6007/IJARBSS/v12-i6/13259
82. Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
83. Woo, D. J., Guo, K., & Susanto, H. (2023). Cases of EFL secondary students'prompt engineering pathways to complete a writing task with ChatGPT. arXiv. https://arxiv.org/abs/2307.05493
84. Woo, D. J., Yu, Y., & Guo, K. (2025). Exploring EFL secondary students'AI-generated text editing while composition writing. 1-31 arXiv. https://arxiv.org/abs/2505.17041
85. Wu, J., Xiao, J., & Nazli, S. (2013). A study of students'assessment in writing skills of the English language. *Educational and Instructional Studies in the World, 6*(2), 129-144. https://dergipark.org.tr/en/pub/eiji/issue/5137/70005
86. Xuan, W. W. (2017). An exploratory study of ESL writing by junior secondary students in China: Text type, register and textual features. *Functional Linguistics, 4*(3), 1-19. https://doi.org/10.1186/s40554-016-0038-9
87. Yadav, AS., Mannow, NA. (2025) Academic anxiety higher secondary school students: urban and ruler areas. *International Journal for Multidisciplinary Research, 7*(3), 1-5.
88. Zeidner, M. *Text anxiety: The state of the art. New York*, NY: Plenum (1998).

How to cite this article: Asmita Ghosh, Surajit Mahato, Subir Sen. Cluster Analysis of Secondary School Students' English Writing Proficiency in Bengali-Medium Schools in West Bengal. *International Journal of Research and Review*. 2026; 13(1): 134-151.
DOI: *https://doi.org/10.52403/ijrr.20260114*