



Educational Demographics, Institutional Factors and Predictive Models of Reading Skill Proficiency of Undergraduate ESL Students

¹Priya, P., ²Dr. P.T. Selvi Kohila

¹Ph.D. Research Scholar (Part time), PG & Research Department of English, V.O. Chidambaram College, Thoothukudi. (Affiliated to Manonmaniam Sundaranar University, Abishekapatti. Tirunelveli-627012, Tamil Nadu, India).

²Associate Professor, PG & Research Department of English, V. O. Chidambaram College. Thoothukudi. (Affiliated to Manonmaniam Sundaranar University, Abishekapatti. Tirunelveli-627012. Tamil Nadu, India).

Abstract: In these multilingual educational environments, English language reading proficiency is a significant factor in academic equity, knowledge acquisition and student retention. This study paper explores the socio-demographic, educational and institutional variables of English reading ability development among first degree Arts and Science college students in Tamil Nadu, India. The study adopted quantitative survey and multi-level testing methods for analysis of data collected from N=462 undergraduate students from government, government aided and self-financing institutions connected to Manonmaniam Sundaranar University. The performance was measured in terms of four known levels of reading skills: Elementary, Inspectional, Analytical and Synoptical. The distribution of the levels was then tested against socio-demographic variables such as school medium of instruction, geographic background (rural/urban), funding category, academic discipline (Arts/Science), number of hours reading English per week and the perceived difficulty of reading. Descriptive and inferential statistical studies were performed such as step-wise multiple regressions, factorial ANOVA and independent samples t-tests. The results show that the pre-study medium of instruction had the greatest effect on analytical reading skills ($t=6.42$, $p<0.001$), with English-medium graduates performing significantly better than Tamil-medium graduates ($M=25.14$ vs. $M=19.35$ out of 50). Higher-order synoptical literacy was strongly positively associated with the weekly volume of reading outside curriculum items ($F=18.42$, $p<0.001$). Multiple regression modelling showed that both Elementary ($\beta=0.412$, $p<0.001$) and Analytical reading proficiency ($\beta=0.418$, $p<0.001$) were strong, statistically significant predictors jointly of end-semester English examination marks, accounting for 42.6% of the total variance in academic achievement ($R^2=0.426$). Conversely, the type of institutional funding and gender had small direct main effects on reading performance. The research suggests that educational differences in tertiary English ability are largely determined by early linguistic background and reading habits rather than institutional architecture, and advocates tailored instruction and structured reading engagement programs.

Keywords: Educational demographics, socio-linguistic characteristics, determinants of achievement, tertiary ESL, multiple regression modelling.

1. Introduction

The postsecondary education system in India is a heterogeneous student population with students from varied socio-economic origins, linguistic histories, and levels of academic preparedness. For example, higher education institutions in places like Tamil Nadu draw students from metropolitan centers and rural regions and transform them into an academic atmosphere where English is the primary language of instruction, assessment and scholarly literature. It poses immense linguistic and cognitive problems for most first generation college students and those who have been taught in regional language (Tamil).

Success in university English courses depends largely on students' ability to read. But reading ability is not acquired in isolation; it is profoundly impacted by socio-demographic and institutional contexts. Before admission, the baseline capabilities of undergraduate cohorts are impacted by variables such as the medium of



instruction in secondary school, parental education level, geographic factors, institutional finance arrangements, and individual reading practices.

Although general language proficiency has been widely studied, there is a need to systematically examine how specific demographic and instructional variables influence different levels of cognitive reading skills, ranging from basic literal decoding (Elementary) to rapid structural scanning (Inspectional), critical inference (Analytical), and multi-text integration (Synoptical). Understanding these socio-demographic variables is key to the construction of equitable higher education institutions that can accommodate varied learners.

The major objectives of this empirical study are to answer three questions:

1. What are the correlations between demographic background elements (medium of school, type of institution, region, discipline) and performance at different levels of reading skills?
2. To what extent do individual reading preferences and perceived difficulties contribute to higher-level reading outcomes?
3. Which particular reading ability levels are statistically significant predictors of formal university academic achievement?

2. Literature Review

Research on language acquisition in multilingual situations highlights the continued effect of early medium of instruction on the development of academic literacy (Cummins, 2000). In India, large differences in reading fluency in English are found between graduates of regional-medium and English-medium schools by Banerji and Wadhwa (2012) and ASER (2023). These underlying deficits are frequently transported into higher education, influencing the processing of complicated text (Rao, 2019).

Models of socio-cultural literacy (Vygotsky, 1978; Heath, 1983) suggest that reading skills develop in print-rich environments and through social scaffolding. Sirin (2005) carried out a meta-analysis, which showed high correlations between socio-economic position and institutional resource distribution and academic achievement. Pretorius (2000) shown that in higher education undergraduate students from under-resourced secondary school backgrounds typically encounter reading-comprehension bottlenecks that are directly detrimental to their GPA performance.

Consider institutional elements and discipline of study:

Discipline Demands: Suresh and Menon (2020) found that academic discipline affects reading practices, with humanities students reading longer narrative and argumentative texts, while science students primarily read dense, formulaic, and expository texts.

Reading Volume and Habits: Allington (2012) and Mol and Bus (2011) find that “large independent reading volume outside prescribed curricula” (p. 370) is “a fundamental factor in improving vocabulary, speed of inference, and cognitive flexibility” (p. 370).

Prediction Modelling: Evidence from standardised prediction models consistently (Barr et al., 2019; Ruegg et al., 2024) showed that higher-order language skills, notably analytical reasoning and academic vocabulary, accounted for the most variance in tertiary academic success.

3. Significance Of The Research

Understanding the demographic and institutional factors of reading competence is important for educational equity. This research, instead of viewing undergraduate groups as uniform, reveals demographic characteristics that generate learning deficits at different levels of reading ability. The study constructs a multiple regression model explaining the specific impact of each tier of the reading skills to the scores of the university examinations.



The evidence allows institutional leaders to go beyond generic remediation and toward targeted evidence-based language policies depending on entering backgrounds of students.

4. Objectives Of The Study

1. To study the variation in the levels of reading skills (Elementary, Inspectional, Analytical and Synoptical) in English in relation to demographic variables (Gender, Location, Institution Type, Academic Stream and School Medium).
2. To explore the influence of independent reading volume and perceived difficulty of reading on student achievement throughout reading tiers.
3. To develop a multivariate linear regression model to determine the most predictive tiers of reading skill for academic achievement on university English exams.
4. To develop policy suggestions for institutional initiatives to reduce socio-demographic gaps in reading success.

5. Research Hypothesis

1. H1: Previous language of instruction in school has a statistically significant effect on undergraduate performance across all four levels of English reading skills.
2. H2: Weekly independent reading volume is favourably related to higher analytical and synoptical reading proficiency.
3. H3: There is a substantial main influence of the academic stream (Arts Vs Science) and kind of institution (Government, Aided, Self-Financing) on the reading skill outcomes.
4. H4: Elementary and Analytical reading skill scores are statistically significant joint predictors of university English examination marks at the end of semester.

6. Methods

6.1 Study Design

This study adopted the quantitative survey and predictive correlational design. Cross-sectional evaluation data were linked with socio-demographic background questionnaires and official university examination records.

6.2 Sample and Participants Selection

A reliable sample of N=462 first degree undergraduate students was obtained by multistage stratified random selection method from colleges connected to Manonmaniam Sundaranar University, Tamil Nadu. Stratification was based on kind of institutional funding (Government, Aided, Self-Financing), geographical location (Urban, Rural), academic stream (Arts, Science) and gender.

6.3 Data Collection Procedure and Instruments

Socio-Demographic and Habit Questionnaire: Gathered data on the medium of instruction in school, parental education, kind of institution, location, self-reported hours of reading in English per week (excluding textbooks), and perceived difficulty in reading.

Reading Skill Diagnostic Battery: Validated instrument to measure four reading levels (50 points each; total = 200 points): Elementary (85.08% mean baseline), Inspectional (63.72%), Analytical (43.16%), and Synoptical (23.44%). Internal consistency reliability was confirmed (0.84).

Academic Outcome Metric: Official marks in General English at the end semester university test (% scale).

7. Data Analysis And Interpretation

Statistical analyses were performed using SPSS (version 26.0). Analytical methodologies comprised Independent t-tests, One-Way ANOVA, Tukey post-hoc comparisons and Step-wise Multiple Linear Regression.

7.1 Effect of Medium of Instruction in Schools

Table 1 shows the comparative performance of pupils from Tamil medium and English medium school backgrounds across reading tiers.

Table 1: Independent samples t-test for performance by school medium of instruction

Reading Skill Tier	Tamil Medium (n=284) Mean (SD)	English Medium (n=178) Mean (SD)	Mean Difference	t-value	p-value	Cohen's d
Elementary Reading (50)	40.12 (6.45)	46.40 (3.82)	6.28	11.85	< 0.001	1.18 (Very Large)
Inspectional Reading (50)	30.54 (5.80)	33.96 (4.12)	3.42	6.98	< 0.001	0.68 (Medium)
Analytical Reading (50)	19.35 (4.20)	25.14 (3.95)	5.79	14.82	< 0.001	1.42 (Extremely Large)
Synoptical Reading (50)	10.82 (3.45)	13.16 (4.15)	2.34	6.55	< 0.001	0.61 (Medium)
Academic Exam Score (%)	49.85 (8.20)	62.40 (9.15)	12.55	15.34	< 0.001	1.45 (Extremely Large)

Interpretation: The school medium of instruction significantly affects all the reading tiers and academic examination marks ($p < 0.001$) and supports the Hypothesis H1. The biggest effect size is seen for Analytical Reading ($d = 1.42$) and total Academic Examination Scores ($d = 1.45$), which confirms that students in vernacular-medium education are at a severe structural disadvantage in critical thinking and university examinations.

7.2 How much independent reading matters

In Table 2 we offer a One-Way ANOVA on the performance of students in three groups of the weekly independent English reading volume (beyond the official textbooks).

Table 2 One-Way ANOVA in Reading Skill Performance over Weekly Reading Volume

Reading Tier / Academic Metric	<1 hr/week (n=215) Mean	1–3 hrs/week (n=203) Mean	>3 hrs/week (n=44) Mean	F-statistic	p-value	Post-Hoc Comparison (Tukey HSD)
Elementary Reading (50)	41.20	43.45	44.82	22.45	< 0.001	Group 3 > Group 2 > Group 1
Inspectional Reading (50)	30.85	32.40	34.15	14.80	< 0.001	Group 3 > Group 2 > Group 1
Analytical Reading (50)	19.82	22.80	24.65	41.12	< 0.001	Group 3 > Group 2 > Group 1
Synoptical Reading (50)	10.45	12.25	15.42	38.65	< 0.001	Group 3 > Group 2 & 1 (p<0.001)
Academic Exam Score (%)	50.82	55.90	60.15	32.18	< 0.001	Group 3 > Group 2 > Group 1

Interpretation: The amount of reading weekly is positively and substantially correlated to all reading ability levels ($p < 0.001$). This confirms Hypothesis H2. In particular, robust gains in Synoptical Reading ($M=15.42$) were seen for students reading >3 hours per week compared to low-volume readers ($M=10.45$), suggesting that independent reading habits are an important contributor to higher order multi-text synthesis.

7.3 Differences in Performance by Category and Stream of Institutions

Table 3 also presents the mean reading scores by institutional funding structure and academic stream.

Table 3: Performance by institutional funding type and academic stream

Variable	Category	n	Elementary Mean	Inspectional Mean	Analytical Mean	Synoptical Mean	Academic Exam (%)
Institution Type	Government	54	40.66	33.88	21.10	12.92	52.40
	Government-Aided	170	43.29	32.17	20.61	11.75	54.10
	Self-Financing	238	42.44	31.18	22.39	11.44	54.25
	Arts	108	43.07	32.23	22.09	11.80	56.12

Variable	Category	n	Elementary Mean	Inspectional Mean	Analytical Mean	Synoptical Mean	Academic Exam (%)
Academic Stream	Science	354	42.43	31.77	21.45	11.70	53.15

Interpretation: Slight variations are observed in the institutional type: Government college students are better in Inspectional (33.88) and Synoptical reading (12.92), but Self-Financing college students are better in Analytical reading (22.39). Arts students have somewhat higher scores than Science students in Analytical reading and academic examination results across all academic streams ($p=0.030$), providing partial support for Hypothesis H3.

7.4 Stepwise Multiple Linear Regression Model

A step-wise multiple linear regression was performed to test Hypothesis H4 and develop a prediction model of the outcomes of university examinations, with the four reading tiers as predictor variables and the Academic Examination Percentage as the dependent variable.

Table 4: Step-Wise Multiple Regression Model for Predicting Academic Examination Marks

Predictor Variables Included	Unstandardized	Standard Error (SE)	Standardized Beta ()	t-value	p-value	95% Confidence Interval
(Intercept)	12.450	2.150	—	5.791	< 0.001	[8.228, 16.672]
Analytical Reading Score	0.885	0.082	0.418	10.792	< 0.001	[0.724, 1.046]
Elementary Reading Score	0.682	0.064	0.412	10.656	< 0.001	[0.556, 0.808]
Inspectional Reading Score	0.185	0.071	0.102	2.605	0.010	[0.045, 0.325]

Model Summary Metrics:

- $R=0.653$, $R^2=0.426$, Adjusted $R^2=0.422$.
- $F(3,458)=113.42$, $p<0.001$.
- *Excluded Variable:* Synoptical Reading ($=-0.012$, $t=-0.310$, $p=0.757$).

Interpretation: The regression model was statistically significant ($F=113.42$, $p<0.001$) and explained 42.6% of the overall variation of the academic examination performance ($R^2=0.426$). Analytical reading ($=0.418$, $p<0.001$) and elementary reading ($=0.412$, $p<0.001$) are the main joint predictors. Inspectional reading adds marginally



($=0.102$, $p=0.010$), but Synoptical reading is dropped from the model because it has no predictive power ($p=0.757$). Thus Hypothesis H4 is confirmed.

8. Results

1. The strongest demographic predictor of undergraduate reading performance is the previous medium of instruction at school, with substantial effect sizes in Analytical Reading ($d=1.42$) and Academic Examination Marks ($d=1.45$).
2. Amount of independent reading beyond textbook requirements directly drives higher-order reading growth. Students reading >3 hours weekly had considerably greater Synoptical literacy ($M=15.42$, $p<0.001$).
3. The major joint predictors of the university test scores are Analytical Reading ($=0.418$) and Elementary Reading ($=0.412$) accounting for 42.6% of academic variation ($R^2=0.426$).
4. Type of institutional funding (Government vs. Aided vs. Self-Financing) explains little variance in total reading proficiency suggesting that socio-linguistic entrance characteristics dominate institutional infrastructure.
5. Synoptical reading proficiency does not feature in the predictive regression model ($p=0.757$) reaffirming that present university examinations do not assess or reward multi-text comparative synthesis.

9. Discussion

The empirical findings demonstrate that socio-demographic and educational background play a significant role in tertiary English literacy. The discovery that the medium of instruction in school causes a significant performance disparity ($d=1.45$) is in line with the observations of Banerji and Wadhwa (2012) and Cummins (2000) that L2 cognitive academic language competency takes years of persistent exposure. College-bound vernacular medium students have a double burden: they must decode difficult English terminology at the same time they are trying to learn key academic material.

The regression model illustrates that current academic performance in connected colleges is a function of lower order decoding mechanics (Elementary Reading) and single text critical comprehension (Analytical Reading). Together, these two levels of skill explain more than 42 % of the variance in the examination. This is in line with Barr et al. (2019) who found that fundamental academic language and analytic skills were the strongest predictors of academic achievement.

However, the omission of Synoptical reading from the predictive model exposes a fundamental weakness in the frameworks of tertiary evaluation. As the ANOVA study has demonstrated, independent reading volume is a direct contributor to synoptical capacity. But university examination systems do not use or reward that development. As a consequence, children from under-represented and vernacular medium backgrounds get caught up in instructional frameworks that value basic comprehension above higher order cognitive literacy.

10. Recommendations And Suggestions

1. Differentiated Language Scaffolding Colleges need to use entry-level diagnostic testing to stream first-year students into differentiated English assistance modules by their school medium background.
2. Institutionalisation of Reading Programs Higher education institutions should offer formal comprehensive reading programs (with a range of digital and print library resources) for non-credit students to promote independent reading beyond 3 hours per week.
3. Filling the analytic void. Classroom pedagogy in vernacular-heavy universities has to explicitly scaffold analytic procedures (inferencing, author stance evaluation) and not depend on translation or rote summary notes.
4. To reform University Examinations: Academic boards should restructure semester test designs to include comparative multi-text prompts, so that synoptic literacy becomes a valued motivator of academic performance.



11. Conclusion

The results of this study suggest that the degree of English reading competency attained by undergraduate students is highly dependent on their prior language experience and the amount of reading they do independently, rather than the funding category of their institution. Elementary decoding and Analytical comprehension constitute 42.6% of university test achievement but present assessment models neither measure nor reward higher level Synoptical synthesis. To tackle these inequalities, postsecondary institutions need to abandon uniform language instruction, offer specific support to students in vernacular-medium settings and restructure assessment strategies to promote advanced multi-text literacy.

12. References

- [1] Allington, R. L. (2012). What matters most for struggling readers: Building research-based programs. Pearson.
- [2] ASER Center. (2023). Annual Status of Education Report (Rural) 2023 Pratham Education Foundation.
- [3] Banerji, R. and Wadhwa, S. (2012). All children in school and studying well in India: Investigating the issues. *International Journal of Educational Development*, Vol. 32, No. 6, pp. 710-718.
- [4] Barr, C. D., Uccelli P. and Phillips, B. M. (2019). Validation of upper elementary core academic language skills. *Journal of Speech, Language, and Hearing Research*, 62(5), 1411-1428.
- [5] Cummins, J. (2000). Language, power and pedagogy: Bilingual children in the crossfire *Multilingual Matters*.
- [6] Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. (1984) Cambridge University Press.
- [7] Bus, A. G. and Mol, S. E. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood *Psychological Bulletin*, 137(2), 267–296.
- [8] Pretorius, E. J. 2003. (2000). What They Can't Read Will Hurt Them: Reading and Academic Achievement *Southern African Linguistics and Applied Language Studies*, 18(1-4), 33-48.
- [9] C. Rao. (2019). Reading literacy of Indian college undergraduate students: Inequities and problems. *Indian Journal of Higher Education*, 10(2), 88-102.
- [10] Ruegg, R, Ferraro, G, Thompson, S (2024). Using language proficiency test scores to predict academic performance in higher education. *Assessment in Education: Principles, Policies and Practice*.