

THE IMPACT OF MULTIMODAL LITERACY PRACTICES ON THE ENGLISH WRITING PROFICIENCY OF PAKISTANI UNDERGRADUATE STUDENTS: A CORRELATIONAL ANALYSIS

Faiza Noor,

(fn435797@gmail.com)

S26W-MPHILE-023, MPhil scholar at University of Okara

Dr. Muhammad Ahmad,

(ahmad453@yandex.com)

Visiting Lecturer, Department of English, University of Okara, Okara, Pakistan.

Shakil Ahmed

(aaashakil73@gmail.com)

S26W-MPHILE-007, MPhil scholar at University of Okara

Abstract

This quantitative correlational study aims to systematically investigate the effect of multimodal literacy practices (visual, auditory, digital interactive media) on the academic English writing proficiency of Pakistani undergraduate students. The data was collected from a validated multimodal literacy practices questionnaire and analyzed based on a standardized IELTS Task 2 academic writing rubric with a random sample of 365 students from various disciplines. Using SPSS, a highly significant and positive relationship was found between multimodal engagement and writing proficiency ($r = .68, p < .01$). Moreover, the results showed that there is a relationship between the application of multiple linear regression between the writing score, and multimodal practices, where the coefficient of determination (R^2) is equal to 0.48. In particular, interactive digital reading was the most significant variable for the improvement of lexical resource, while visual/video media significantly boosted the structural coherence. Key results show that the incorporation of students' current digital multimodal practices into official ESL teaching programs can be a strategy to overcome gaps in system and can significantly impact the curriculum design and pedagogical reform in the Pakistani higher education sector by providing valuable evidence-based insights to curriculum developers.

Keywords: academic English writing; digital interaction; ESL learners; higher education in Pakistan; multimodal literacy; pedagogy of multiliteracies

1. Introduction

1.1 Background of the Study

Literacy is more than just the ability to read and write printed words in today's digital 21st century and has become a complex, multifaceted framework that is referred to as multimodal literacy. This paradigm is a dynamic interplay between text, audio, spatial, gestural and visual ways of meaning making, which in turn changes the way people understand, process and create meaning (Mills, Unsworth & Scholes, 2023). In today's global academic landscape, which is rapidly digitizing, being familiar with these different modalities, and knowing how to use them is no longer a second language but a first one. This global movement is significant in relation to the English as a Second Language (ESL) teaching and learning in the Higher Education Institutions in Pakistan.

Traditional, print based approaches have been the base of the Pakistani academic world over a period of decades in the field of language learning. But the access of the internet, smartphones and digital media has also become an important part of the life of undergraduate students in Pakistan which has resulted in a significant change in their linguistic practices (Ameen, Saleem, & Taj, 2023). As a result, there is a reorientation of traditional conceptions of literacy towards a range of multiliteracies. This deep change not only influences passive language input, which is what students learn by listening to a variety of digital resources including podcasts, educational videos and interactive online resources, but also how they cognitively represent language and actively use it. The use of multimodal texts in these students' everyday lives is driving their cognitive linguistic processing, and to make full use of

their digital practices for formal improvement in their academic writing, pedagogical strategies need to change along the same lines.

1.2 Problem Statement

Although the modern student body is highly immersed in the digital world, Pakistani undergraduates still have deep and longstanding problems in academic English writing. Such a problem often presents itself as either a constant failure to express complex thoughts with clarity, a small range of vocabulary, or frequent errors in grammar. The systemic deficiencies are largely caused by the fact that the country still sticks to traditional approaches of grammar translation and rote-learning inculcated in the system at the ground level (Fareed, Ashraf, & Bilal, 2016). This highlights a pedagogical gap, that is, students are consuming, navigating and creating rich digital and multimodal texts in their everyday lives, but this is something that is not being utilized in the context of higher education in Pakistan. Now to fill this gap, the present study investigates the statistical relationship and predictive power of multimodal literacy practices on the academic English writing ability of Pakistani undergraduate students to bridge the gap between students' current digital consumption and their formal ESL academic writing outcomes.

1.3 Scope and Delimitations of the Study

This study was undertaken within highly specific, predetermined boundaries, in order to be methodological and to allow for targeted data analysis. This research only aims to assess the English academic writing skills and digital multimodal habits of undergraduate students currently studying in Higher Education Commission (HEC) recognized universities in Pakistan. The study is conducted in three major academic disciplines, namely Arts & Humanities, Natural & Applied Sciences and Engineering & Technology to have diverse representation.

In terms of delimitations, the independent variable (multimodal literacy practices) is narrowed down to digital multimodality, meaning visual/video media, digital interactive reading and auditory reading in the form of podcasts, but not physical or face-to-face multimodal practices (such as gestures in live performance or physical art making). Moreover, the dependent variable is well-defined as formal academic writing skills which are assessed using a standard IELTS Task 2 argumentative writing rubric. The study specifically does not include other areas of language acquisition (e.g., speed of reading, listening or spoken fluency) to narrowly focus on the most important academic product demanded from university students. Lastly, the study is limited to a cross-sectional time period, which means that it only reflects the current situation and not the developmental changes over time.

1.4 Research Objectives

This study seeks to achieve three primary academic objectives:

1. To profile Pakistani undergraduates' everyday multimodal literacy practices and benchmark their academic English writing proficiency using a standardized global rubric.
2. To evaluate the statistical correlation—including strength, significance, and direction—between students' multimodal literacy practices and their writing proficiency scores.
3. To determine the predictive effect of specific multimodal inputs (visual, auditory, interactive) on localized writing dimensions (lexical resource, structural coherence, and grammatical range).

To conduct an in-depth study of the kinds of multimodal literacy activities (from the viewing of audio-visual media to the practice of digital interactive reading) that Pakistani undergraduate students are doing in their everyday lives and benchmarking

of the current English academic writing level of such students on a world-wide recognized and standardized assessment rubric;

To empirically identify the strength, statistical significance and direction of the correlation between multimodal literacy practices and the multimodal literacy English writing proficiency scores of students.

Deeply analyse the predictive effect of specific multimodal inputs (e.g., visual, auditory, digital interactive) for specific localized aspects of writing: analysing lexical resource, structural coherence and the range of grammatical knowledge.

1.5 Research Questions & Hypotheses

1. What are the current multimodal literacy practices of Pakistani undergraduate students, and how do they perform in academic English writing when assessed against a standardized rubric?
2. What is the statistical relationship between the frequency of students' multimodal literacy practices and their level of academic English writing proficiency?
3. To what extent do specific multimodal inputs (visual, auditory, and digital interactive) predict distinct dimensions of academic writing, namely lexical resource, coherence, and grammatical range?

This study has primary research questions and testable hypotheses to directly respond to the research objectives:

1. What are the foundational levels of multimodal literacy practices and performances of participants? The level of academic writing in English of the Pakistani undergraduate students?
2. How is the frequency of the variable x statistically related to y ? The range of multimodal literacy practices and the degree of academic English writing. In the designated population, do they have a good command of the language?
3. Which specific multimodal practices (visual, auditory, digital etc.) are useful in the classroom? Two variables (hyperactivity and interactive) are highly predictive of unique aspects of academic writing Lexical resource, coherence and grammatical range of proficiency?

Drawing on input-output theories in second language acquisition (e.g., Krashen's input hypothesis (1985), Swain's (2005) output hypothesis), the following directional hypotheses were formulated for quantitative testing:

- **Alternative Hypothesis (H1):** There is a statistically significant positive correlation between the level of multimodal literacy practices and English writing proficiency scores of the Pakistani undergraduate students with ($p < .05$).
- **Null Hypothesis (H0):** There is no statistically significant relationship between multimodal literacy practices and English writing proficiency of the target population, which means that the use of digital multimodal content is not statistically significant to formal academic writing outcomes.

1.6 Significance of the Study

This research is significant at two levels: theoretical and practical. This research has a strong theoretical contribution to the vast field of pedagogy of multiliteracies as it makes an empirical quantitative attempt to validate the idea in the under-researched South Asian ESL context. It is not an anecdotal evidence and instead offers statistical facts about the translation of multimodal inputs to cognitive linguistic gains. Practically, this study is important for institutional actors. That is, the results of this study will be used as a ground for formal inclusion of digital and multimodal tasks specifically in formalized university syllabus by curriculum developers and Higher Education Commission policymakers in Pakistan. Moreover, the research will shed light on how to develop specific strategies for ESL teachers to effectively

employ students' current practices (e.g., using interactive texts, listening and viewing audio-visual material) to learn and address specific ESL writing deficits and to modernize the face-to-face approach to teaching writing nationwide.

2. Literature Review

2.1 Multimodal Literacy

This study is based on the theoretical context of multimodal literacy which questions the traditional, text-based conception of reading and writing. This change can be attributed to the ground-breaking approach to literacy of the New London Group (1996) called the 'pedagogy of multiliteracies'. Being aware of the significant changes in the global communications environment, they felt that the pedagogy of literacy needed to be far-reaching in its consideration of the increasing diversity of text forms that is linked to contemporary information and multimedia technologies (Tweedale, 2025). This basic framework was later enriched by Cope and Kalantzis (2016) who placed emphasis on the process of meaning making in the 21st century being not necessarily limited to one isolated mode. Rather, it is the interplay of multiple (linguistic, visual, auditory, spatial and gestural) elements. Thus, in the present 'reading' other scriptures like alphabetic ones have to be synthesized with moving images, and interactive hypertexts in order to be understood (Cope & Kalantzis, 2016).

This evolution calls for a clear shift in the way that text is consumed from monomodal (print-only) to multimodal forms of engagement, which has been extensively theorized by Keane (2025). Multimodality, as Keane suggests, is very important to contemporary communication in that affordances (specific cognitive potentials and limitations) are present in each specific mode of communication and are used in the process of meaning-making (Keane, 2025). For example, written text is sequenced, and follows the logic of time; the visual modes are spatial and follow the logic of simultaneousness. For ESL learners, when they are reading a multimodal text, like a documentary with sub-titles or an interactive digital text, they are not just interpreting alphabetical words, they are also inferring the visual elements, the spatial organization and the auditory tones to create a whole language understanding. Multimodal literacy is at the core of this multi-faceted cognitive processing. Moreover, when these various semiotic materials are integrated, the learners are able to cross-check the meaning if they do not understand a language element, they can take it up from the other semiotic materials. It is therefore of crucial importance when considering the modern ESL learner, who has almost exclusively multimodal digital experiences of the English language outside the classroom (Park, 2023).

2.2 English Writing Proficiency in the Pakistani Context

The situation of English writing in Pakistan is very critical and it is a source of concern. It is essential to discuss the current situation of English writing achievement in the context of Pakistan education system and in this regard it is important to critically address the need to investigate multimodal literacy. In Pakistan, language teaching has traditionally been grammar-based, teacher centered and predominantly the grammar translation method has been in use (Ahmed, 2025). In the traditional and highly formalized settings, the pedagogical concern is limited to the rote memorization of the rules of grammar, memorization of list of words out of context and mechanical translation of individual sentences from Urdu to English. The traditional mono-modal and print-based teaching and learning of a language essentially deprives the language of its communicative nature and makes the students quite inadequate to cope with the communicative dynamic and higher order demands of academic discourse in university environment.

Literature shows the serious concern about the out-dated models of pedagogy. For example, Fareed, Ashraf and Bilal (2016) found that undergraduates often need a very limited range of lexis, and use basic and repetitive words that do not convey the complexities of

academic language. In addition, they tend to lack coherence in the structure of their writing; the students have great difficulties in arranging their ideas logically and make their paragraphs incoherent and the lack of micro and macro cohesion is severe. These results have been well supported by more recent empirical studies conducted by Nazar and Nordin, (2024), which highlight the lack of task achievement and the high rate of the grammatical errors in complex sentence constructions. Students do not receive a multimodal input that contextualizes grammar and vocabulary in various ways, so that their written product is affected, artificial and formulaic (Nazar & Nordin, 2024). They are taught theory which is theoretical knowledge of the language, not how to use the language in a range of meaningful contexts. These problems suggest a failure of existing curriculum to meaningfully engage learners. Various pedagogical interventions have been proposed to address these writing deficiencies. For instance, Leotta and Ahmad (2025) demonstrated that explicit instruction combined with process writing strategies can yield measurable improvements in academic English writing skills among learners. While such teacher-directed, classroom-based approaches offer structured support, they remain largely monomodal and text-centric, often overlooking the rich, digitally mediated multimodal input that students naturally engage with outside formal instructional settings. The present study argues that rather than replacing explicit instruction, multimodal literacy practices can serve as a powerful complementary scaffold (i.e., the one that influences students' existing digital habits to enhance lexical resource, coherence, and grammatical range). If Pakistani students are not exposed to the natural flow, tone and diverse structures of English, which are readily available with multimodal mediums, they will continue to be affected with these writing shortcomings and it becomes an absolute necessity of academic education to explore digitally-driven sphere of literacy practices (Khan, Mushtaq, Bakhsh, Butt & Khan, 2025). Empirical corpus-based studies confirm these persistent difficulties. Ahmad, Mahmood, and Siddique (2019) documented significant deficiencies in organizational skills and cohesive devices among Pakistani academic writers. More recently, Ahmad, Mahmood, and Siddique (2025) found that even doctoral students exhibit measurable developmental gaps in L2 academic writing, suggesting that traditional pedagogical approaches fail to address foundational writing issues across educational levels.

2.3 The Intersection of Multimodal Practices and ESL Writing

The use of multimodal literacies in relation to ESL academic writing is a very promising, but complex aspect in applied linguistics. Research from around the world has started to explore how multimodal inputs in the classroom positively affect second language (L2) writing. The concept of multimodal (allowing for a variety of ways of viewing, listening and engaging with educational content in the classroom, such as through watching educational videos with subtitles, engaging in digital readings that include interactive components, or listening to academic podcasts) is often cited as a powerful method of reducing the affective filter (Das & Singh, 2023). Multimodal texts can present complex linguistic information in entertaining, non-threatening ways, making assimilation of complex grammar and new words much less threatening. Anderson, Anson, Townsend, and Yancey (2006) provided preliminary research that showed that when students learn and are engaged in multimodal writing in college composition courses, they are more motivated to learn and deep cognitive retention is enhanced. If students can hear the language and read the text (exact subtitle) at the same time, and see the exact same image on screen, the neurological memory tracts will be reinforced much more than by reading the plain text (Tan, 2023).

Most importantly, it is a rich multimodal exposure that will help develop vocabulary and will naturally contextualize grammar, resulting in better academic writing skills. If students are interacting with an interactive article or a video essay, the language they are reading or viewing is in the authentic context that the language is used in the real world. They implicitly

notice the way that native or fluent speakers use linking words to achieve text cohesion, how they use lexical diversity to keep the audience engaged and how they use complex structures to effectively link them together to make a nuanced point. A powerful and internalized cognitive template can be achieved by this multi-sensory observation and then used by students when asked to write their academic papers. For instance, learning to arrange information in an infographic in a logical way could help students learn how to categorize information logically, that is, a cognitive skill which can be easily transferred to paragraph structuring and coherence in an argumentative essay (Zhang & Yu, 2023). Likewise, by listening to podcasts regularly, a student can naturally improve his or her command of sentence structure, intonation and sentence pacing, and thus produce more fluent and precise sentences in the written language. Multimodal texts are a kind of invisible and very powerful ESL tutor because it connects with the students' passive language experience and helps them to produce language. They allow for the sustained rich language modelling that is missing in the textbook writing exercises of Pakistani classrooms, and see multimodality as a dynamic and transformative tool in writing development.

2.4 Gaps in Existing Literature

The review indicates that there are significant gaps in the existing literature. A theoretical rationale underlies the promotion of multimodal literacy, but there is a significant absence of empirical research. First, the vast majority of multimodal research (e.g., Anson, Dannels & St. Clair, 2005; Das & Singh, 2023; Hicks, 2014; Lim & Tan-Chia, 2022; Serafini, 2015; Tan, 2023; Tweedale, 2025; Unsworth, Cope & Nicholls, 2019; Yi, 2014; Zhang & Yu, 2023; Walsh, 2010) has been conducted in the countries other than Pakistan where there are very advanced digital infrastructure and student-centred curricula. The results of the study, therefore, have limited applications to the unique socio-cultural and educational context of South Asia, particularly Pakistan. Second, and most importantly, the existing literature mostly consists of qualitative studies. Past research has tended to be based on student interviews, which are subjective, and case studies of observations. Only a few studies, and none of them are systematic or quantitative, have measured the direct, statistical effect of multimodal practices on standardized academic writing measures (e.g., lexical resource, grammatical accuracy, and coherence). This study provides a critical solution to this empirical gap by moving from qualitative assumptions to statistical validation and presents conclusive statistical data to back up pedagogical changes in the Pakistani context of ESL.

3. Methodology

3.1 Research Design

The quantitative explanatory correlational design has been utilized in this study because it is able to rigorously answer the research questions proposed and adequately test the formulated research hypotheses. Thomas and Zubkov (2023) state that a correlational research design is only used when a researcher wants to find associations, statistical relationships, and predict variance between two or more variables without artificially manipulating these variables in a controlled/experimental environment. The current study's focus is not to experiment with students' existing and naturally occurring digital habits, as it is impractical and ecologically invalid. Thus, the use of explanatory correlational design will be the appropriate design in systematically measuring independent variable (multimodal literacy practices) and dependent variable (English writing proficiency) which are naturally at hand in the target population. In doing so, this quantitative approach, which is not experimental, effectively provides a precise, step-by-step picture of the exact relationship and the extent of the relationship between current patterns of digital consumption and the formal academic outcomes. In addition, the explanatory nature of the design means that the researcher can go on to make future predictions, which would enable them to account for variations of the

standardized writing scores in terms of the different types of multimodal literacy. This would result in extremely objective and generalizable empirical data.

3.2 Population and Sampling

The entire study was oriented towards the population of undergraduate students in public and private universities (Higher Education Commission) recognized in Pakistan. Given the diversity of socio-economic and academic disciplines of the Pakistani higher education context, the need for a highly structured sampling technique was essential to ensure a broad representativeness and deal with selection bias. Hence, sampling was done by Stratified Random sampling technique. The population was carefully segmented into three major academic categories (i.e., Arts and Humanities, Natural and Applied Sciences and Engineering and Technology). Random selections were then made within each separate division to make up the final sample, again making sure that the various cognitive and linguistic demands of academic disciplines were represented in the data.

The number of participants (N) required for the study was determined before the data collection using G*Power 3.1 software, with the settings of a linear multiple regression analysis with an anticipated medium effect size ($f^2 = 0.15$), an alpha value of .05 and a strict statistical power of 0.95, resulting in a minimum requirement of 280 participants. To account for possible dropouts, misfilled or incomplete questionnaires, and non-responsiveness of survey items, a buffer was added to the study, with a minimum target of 350 to 400 undergraduate students. This is a strong sampling technique as the findings from the sample are very likely to be representative of the entire population of Pakistani undergraduates and reliable enough to draw inferential statistics.

3.3 Instrumentation

Two main instruments were used in collecting data in this study; their validity and reliability were adequately checked to ensure that the variables investigated were quantified accurately.

Multimodal literacy practices questionnaire (MLPQ) - independent variable 1. The MLPQ was used to evaluate the level of students' digital habits. This instrument is a well-structured, self-reported, 5-point Likert scale (from 1 = Never to 5 = Always) that has been carefully adapted from digital literacy scales validated across the world and highly customised according to the socio-educational context of Pakistan. The MLPQ consists of 25 items strategically divided into three different sub-constructs: (a) Visual and Audio media (e.g., watching subtitled documentaries, academic YouTube videos, or listening to English podcasts), (b) Spatial and gestural media (e.g., navigating complex infographics and interactive multimedia presentations), and (c) Digital interactive texts (e.g., participating in English-medium discussion forums, reading hyperlinked digital articles, and engaging with social media discourse).

2nd Instrument: standardized academic writing test (dependent variable) In order to objectively assess the participants' English writing proficiency, they were given a standardized timed writing task with a similar writing prompt as in the International English Language Testing System (IELTS) Academic Task 2. Participants had to write an argumentative essay of 250 words in a time-limited manner of 40 minutes.

The resulting essays were carefully marked with a global standard, internationally recognised marking scheme, to ensure that subjective grading bias was removed. This rubric clearly separates the criteria to be evaluated and assigns a score for each one of the four that are important: Task Response (Formulate and support a clear argument), Coherence and Cohesion (Logical sequence of ideas and use of transitional devices), Lexical Resource (Range, precision, and sophistication of lexical resources), and Grammatical Range and Accuracy (Variety of complex sentence structures and absence of syntactic errors). A standardized band

scale from 1-9 was used to score each criterion, which were then combined to make up the writing proficiency score.

3.4 Data Collection Procedures

A thorough and standardized administrative procedure and ethical guidelines were followed during the data collection process. All participants signed a formal informed consent and prior to any data collection, formal institutional review board (IRB) approval was obtained. Students were well informed about the voluntary nature of the study and that they were to remain absolutely anonymous; all test materials were coded with numbers so that no personal information could be associated with a writing score.

The administration took place in the following two stages. In the first part, the MLPQ was digitally published through Google Forms that enabled a systematic self-report of multimodal habits by students. After the questionnaire, a standardized academic writing test was administered in a well-supervised classroom to completely stop students from cheating with artificial intelligence tools, digital dictionary or other people so that the written products are true. After obtaining essays, a strict double blind grading process was followed. Two trained and certified independent ESL teachers, unaware of the students' MLPQ scores, examined the essays, scoring them according to the above rubric. Since the scoring was done by two graders, Cohen's kappa coefficient was used to measure interrater reliability and to verify consistency between the two graders prior to statistical analysis of the scores.

3.5 Data Analysis

At the end of the data collection period, all the raw quantitative data collected were put into a Statistical Package for the Social Sciences (SPSS version 28) for thorough statistical analysis. However, prior to completion of the main inferential analyses, preliminary data cleaning was done to determine if there were any extreme error values that needed to be removed and Shapiro-Wilk tests were performed to confirm that the data was normally distributed.

The analytical phase was carried out in three different steps. The first step was to determine internal consistency and reliability of the MLPQ in the local context using Cronbach's Alpha coefficient. Second, a Pearson Product-Moment Correlation Coefficient (r) was calculated to directly answer Research Question 1 and test the main hypothesis (H1). This was the final test to determine the magnitude, direction and statistical significance of the overall link between multimodal literacy practices and English writing. Finally, Multiple Linear Regression analysis was performed to test the prediction effect of the independent variables. In the present model, the sub-constructs of the MLPQ (visual, auditory, digital interactive) were analyzed as the predictors and the different domains of writing (lexical resource, coherence etc) were used as the predicted variables. This regression analysis was important to find and identify, which of the multimodal practice is the best statistical predictor of academic writing success among the Pakistani students.

4. Results

4.1 Descriptive Statistics of Multimodal Practices

The descriptive statistical analysis of the MLPQ (Table 1) proved to be critical and detailed in nature as it offered a valuable look-in at the particular digital consumption habits of the Pakistani undergraduate sample ($N=365$). The overall data showed a moderately high degree of overall multimodal engagement with an overall mean of 3.82 ($SD = 0.65$) on the 5-point Likert data. A detailed look at the MLPQ sub-constructs, however, revealed discrepancies among the kinds of modalities that were most preferred among the demographic.

Visual and Video consumption sub-scale had the highest frequency of engagement ($M = 4.25$, $SD = 0.58$). This means that learners in Pakistan overwhelmingly make use of video based websites (English language YouTube tutorials, subtitled digital documentaries and short form

video content) for meaning making and language input. Digital interactive and social was the other sub-scale with a high level of engagement ($M = 3.95$, $SD = 0.72$), with the majority of students engaging with digital reading of hyperlinked articles, blogs, and interactive social reading/writing. The auditory/podcast sub-scale in contrast, had a much lower mean score, $M = 2.45$, $SD = 0.88$. Here it seems that the strictly audio-based multimodal formats, for example in the academic world or audiobooks in English, are very underused, perhaps because of a culture of teaching and learning that is more visual than auditory and perhaps because there is a lack of formal institutional support in the use of such tools.

4.2 Descriptive Statistics of Writing Proficiency

This study presents the general results of students' writing proficiency (Table 1) in this study. Standardized Academic Writing Test assessment was used to get a benchmark assessment of the sample population's English writing skills. On a 9-band scale, students in the demographic achieved an intermediate, but functionally limited, general mean score in writing proficiency of 5.4 ($SD = 0.85$) which has difficulty with the more demanding academic writing tasks.

An itemization in the four standardized grading rubrics was precise and revealed clearly certain linguistic strengths and systemic weaknesses. Participants' scores on Task Response were moderate ($M = 6.1$, $SD = 0.70$), indicating some level of understanding of the requirements of the prompts and the ability to develop a central argument that is relevant to the prompts. But the results were terrible in terms of the more technical and structural aspects of academic writing. The mean score for Coherence and Cohesion was a worryingly low $M = 5.0$, $SD = 0.82$ indicating the students' extreme difficulty in logically ordering ideas, using complex linking words and phrases, and ensuring paragraph cohesion. This finding aligns with Ahmad, Mahmood, and Siddique (2019), who reported that Pakistani academic writers struggle notably with logical sequencing and cohesive devices. The persistence of such issues, also observed at the doctoral level (Ahmad, Mahmood, & Siddique, 2025) highlights the systemic nature of the problem and strengthens the argument for multimodal pedagogical interventions. In a similar way, Grammatical Range and Accuracy was found to be the lowest component overall with a high incidence of systemic errors in complex sentence construction and continued errors with subject-verb agreement. Lexical Resource was at a moderate level ($M = 5.5$, $SD = 0.75$) but the graders commented that there was a great use of basic, repetitive vocabulary instead of considering more advanced, academic collocations. The descriptive findings confirm the negative impact of traditional rote-learning processes that were used in Pakistan's educational past.

Table 1 Descriptive Statistics for Multimodal Literacy Practices and Academic Writing Proficiency ($N = 365$)

Variables and Sub-Constructs	Mean (M)	Std. Deviation (SD)	Min	Max
Multimodal Literacy Practices (IV)	3.82	0.65	1.2	5.0
Visual and Video Consumption	4.25	0.58	2.0	5.0
Digital Interactive and Social Texts	3.95	0.72	1.5	5.0
Auditory / Podcast Consumption	2.45	0.88	1.0	4.5

Academic Writing Proficiency (DV)	5.40	0.85	3.5	8.0
Task Response	6.10	0.70	4.0	8.5
Lexical Resource	5.50	0.75	3.5	8.0
Coherence and Cohesion	5.00	0.82	3.0	7.5
Grammatical Range and Accuracy	4.80	0.95	2.5	7.0

Note. Multimodal practices were measured on a 5-point Likert scale (1 = *Never*, 5 = *Always*). Writing proficiency was measured on a standardized 9-band academic rubric (1 = *Non-user*, 9 = *Expert*).

4.3 Correlational Analysis

A Pearson Product-Moment Correlation Coefficient (Table 2) was performed to test the main Alternative Hypothesis (H1) which stated that there would be a statistically significant positive correlation between multimodal practices in literacy and English writing. The correlation matrix that was generated had very significant results and demonstrated a rejection of the Null Hypothesis (H0).

The overall statistical analysis demonstrated that the students' overall multimodal literacy scores and their overall academic writing proficiency scores were significantly positively correlated with a high value, $r = .68, p < .01$. This strong r-value is clear proof that higher engagement of the Pakistani undergraduate students with different multimodal digital texts, in more frequency and more intensity, leads to a higher and predictable improvement in their academic writing skills in standardized format. In addition, correlational tests were conducted between selected variables in specific areas to reveal important details. A very high positive correlation, for example, was found between the Digital Interactive reading sub-scale and the Lexical Resource writing rubric ($r = .72, p < .01$). This indicates that active text interaction in hyperlinked digital text environments is a key factor to increase a learner's active vocabulary. Furthermore, there was a moderate-to-strong positive relationship between Visual/Video consumption and Coherence and Cohesion ($r = .58, p < .05$), indicating that well-produced digital videos, which have a logical and structured flow of ideas, can facilitate students in internalizing macro-level organizational skills that will be transferred to their essay writing.

4.4 Regression Analysis

To go beyond the level of association and to establish the precise prediction effect of the independent variables, Multiple Linear Regression analysis was used. In this predictive model, the three separate multimodal sub-constructs (Visual/Video, Digital Interactive, and Auditory) were used as predictor variables and was analyzed in combination with the overall academic writing proficiency score as the dependent variable.

The model summary showed that there was a significant R^2 of .48 that means 48% of the total variance of the students' English writing proficiency score was accounted by the students' multimodal literacy practices directly. The ANOVA output was highly statistically significant, confirming that this regression model was highly statistically significant, $F(3, 361) = 45.32, p < .001$.

The standardized Beta (β) coefficients were evaluated, and this identified the specific order of predictive effect. Digital Interactive practices came out as the best positive predictor of writing proficiency, with a higher value (i.e., $\beta = .42, t = 6.85, p < .001$), indicating that active participation in digital reading, social interactions, and online forums has

the greatest positive impact in improving writing skills, especially vocabulary acquisition. There was also a significant moderate predictive impact of Visual/Video media consumption, specifically with regards to structural coherence ($\beta = .29, t = 4.12, p < .01$). In contrast, Auditory/Podcast proved not to be a strong predictor of writing ability ($\beta = .06, t = 1.05, p = .29$). For the auditory mode of input, this result is not statistically significant, which means that, in the present socio-educational context of Pakistan, listening to language without any visual or interactive stimuli does not necessarily produce any significant measurable impact on writing output in educational context.

Table 2 Multiple Linear Regression Predicting Overall English Writing Proficiency from Multimodal Practices

Predictor Variables	Unstandardize d B	Std. Error	Standardize d β	t	p
(Constant)	1.85	0.32	-	5.78	<.001
Digital Interactive Texts	0.45	0.06	.42	6.85	<.001*
Visual and Video Consumption	0.38	0.09	.29	4.12	.004*
Auditory / Podcast Consumption	0.08	0.07	.06	1.05	.290

Note. $R^2 = .48$, Adjusted $R^2 = .47$, $F(3, 361) = 45.32, p < .001$.

- = statistically significant predictor at the $p < .01$ level.

5. Discussion

The strong quantitative results obtained in this study give clear empirical evidence in support of the alternative hypothesis (H1) that there is a highly significant positive correlation between multimodal literacy practices and academic English writing proficiency of Pakistani undergraduate students, with a correlation coefficient (r) of .68. In order to gain a comprehensive understanding of the statistical data that expressed itself in this way one needs to reflect critically the underlying cognitive mechanics of the acquisition of language. The regression analysis showed that the interactive digital reading and visual/video consumption had a significant contribution of 48% to writing scores. This is a powerful effect and can be explained in terms of the concept of cognitive mapping. As L2 learners engage with diverse modes of representation such as a hyperlinked, researched text or a visually arranged documentary that is engaging, their not learning language in a vacuum, but rather processing the language in context, as in space, and as in image.

It's a rich, multidimensional exposure in effect building a very complex cognitive scaffold. For example, well-designed educational videos have a linear order of presentation and a logical flow of information, thus giving students an internalized mental model to follow at the macro level of textual organization. Thus when asked to write an academic essay, students unconsciously borrow from these internalized visual-narrative schemas and the result is an improvement of scores in the coherence and cohesion grading criteria. Likewise, digital interactive reading had a strong predicting effect on the lexical resource score with the coefficient value of $\beta = .42$, which was also a logical one. Digital discourse requires students to engage with the target language in a variety of registers, tones and very context-dependent

situations. This highly dynamic, learner-centered approach to reading and using living, hyperlinked texts radically speeds up the incidental learning of sophisticated academic collocations and diverse syntactic structures, enabling students to have a much wider and more flexible range of vocabulary than rote learning could ever provide.

The results of this present study contrast to those discussed in the wider ESL literature, as well as to other more recent studies. The basic descriptive statistics about the students' writing skills strongly support the basic observations that Fareed, Ashraf, and Bilal (2016) made. Fareed et al. correctly identified the prevalent and systemic writing problems of traditional teaching which include the very limited range of grammatical structures and coherence that can be attributed to the excessive reliance on grammar translation method and rote learning in Pakistan. The mean score for grammar ($M = 4.8$) and cohesion ($M = 5.0$) in the present study substantiate this point on an empirical level, and show that these historical deficiencies are still very much entrenched in the general student body.

The current study however, is unique from historical notions as it demonstrates that students who engage in multimodal literacy independently are remarkably able to overcome these traditional barriers. This particular phenomenon ideally fits with, and is strongly supportive of the theoretical claims of scholars like Tan (2023). Tan claimed that the affective filter can be lowered by the use of digital and multimodal environment and that it can actively engage the higher-order cognitive processing, which can significantly facilitate the production of text and the ideation in L2 learners. The results of regressions in the present study are statistically validations of Yi's (2014) claims in the South Asian context. It demonstrates that the multimodal interaction does not simply entertain but it also builds higher order thinking skills and language mapping for successful academic writing. This research finally brings together the diagnosis of the problem by Fareed et al. (2007) and the multimodal solution by Yi (2014) to firmly establish the need for a radical pedagogical shift in order to address the traditional ESL writing problem in Pakistan, which students can successfully manage in their everyday lives through the digital modalities.

5.1 Pedagogical Implications

The empirical data produced in the present study has implications fully transformative in nature which require the attention of all stakeholders in the institutions, curriculum designers and Higher Education Commission of Pakistan to make them implementable. There is a clear and undeniable statistical correlation and predictive power of multimodal practices that need to be addressed with a comprehensive and systemic change of curriculum. What was once considered to be good pedagogy or acceptable is no longer acceptable: the sole use of static, monomodal, textbook writing prompts that turn off the modern learner. Rather, university syllabi should be revised to include the dynamic multimodal input directly in the process of academic writing. For instance, instead of giving a text-based prewriting task for an argumentative writing assignment, ESL teachers can make use of rich multimodal prewriting tasks. Students might be asked to critically interpret the rhetorical strategies and the way information is presented in complicated visual graphics from an academic TED Talk or be asked to synthesize information from complex digital infographics before drafting an essay. These multimodal texts are used as primary sources, giving students authentic, highly contextualised language models that engage students in connecting their digital consumption with formal academic production.

In addition, such a curriculum change will not be accomplished without targeted, widespread professional development for teachers. Many teachers in the universities of Pakistan were trained in strictly traditional paradigm and might be lacking those particular digital multiliteracies needed to implement this modernized pedagogy. To ensure that ESL teachers are able to assess, curate and use multimodality effectively in their classrooms. It is

therefore important that universities invest in rigorous training workshops to equip ESL teachers with the technical skills and theoretical knowledge. Teachers need to be able to create grading rubrics that assess the student's capacity to create a multimodal synthesis of written research as well as the written product. In the end, the official recognition and use of the multimodal digital tools that students already have, can help in successfully moving from the model of language teaching in higher education towards a model of academic multiliteracies in the 21st century in Pakistan.

6. Conclusion

The main aim of this study is to explore the statistical relationship and predictive power of multimodal literacy practices to academic English writing proficiency of the undergraduate students of Pakistan. Data was obtained from a diverse, stratified sample of 365 undergraduate learners from various disciplines of learning through a highly rigorous quantitative explanatory correlational design. In addition to cross analyzing the self-reported scores on the MLPQ with the scores collected from a double-blind scored Standardized Academic Writing Test, the study resulted in highly significant empirical findings. The results proved that there was a positive correlation between the data, with $r = .68$. Additionally, the multiple linear regression showed that multimodality is a strong predictor of academic achievement, with 48% of the variance of writing scores explained. In particular, it was found that the use of interactive reading texts has the greatest statistical association with increased vocabulary resources, and that viewing and video media use is statistically associated with increased structural coherence. On the other hand, only the strictly auditory inputs showed very low predictive values by highlighting the need for interactive, visual and multifaceted, digital pedagogical interventions in the language teaching of the modern times.

6.1 Limitations of the Study

There are a number of this research that needs to be acknowledged due to the rigour of the methodological framework used. Most significantly, the study is based on one main research approach, which is a correlational quantitative design; even though there is a strong statistical predictive relationship, this does not necessarily indicate absolute causation. We are confident in saying that multimodal practices and higher writing scores go hand in hand, but we cannot say that with a high degree of certainty without experimental manipulation. Furthermore, a regional bias may have arisen as the sampling was stratified random sampling, but the sample was mostly concentrated in the big city education hubs of Punjab and Sindh. Therefore, the results should not be directly generalised to regions of Pakistan with restricted technology access, where it will have less impact on students' learning. Last but not least, because of the cross-sectional nature of the data collection, it only reflects a single moment of proficiency, and does not reflect the longitudinal and developmental trajectory of students' writing skills as their multimodal habits develop over a longer academic timeframe.

6.2 Recommendations for Future Research

Several possibilities for future research are recommended as there are structural limitations associated with this study and the empirical data provided by this research needs to be built upon. Future studies should focus on developing research designs that are highly controlled and experimental or quasi-experimental, with pre-test/post-test control groups, to prove that there is a direct causal link between the specific multimodal pedagogical interventions and measurable gains in academic writing metrics. Secondly, to see the whole trajectory of English language learning, it would be interesting to see the longitudinal changes in the writing abilities of one group of students over the four years of their undergraduate studies. Finally, the quantitative framework should be very much complemented by solid qualitative follow up. In-depth phenomenological student interviews or the use of 'think aloud' cognitive protocols would enable researchers to explore in depth the internal mechanisms of

how learners process, translate and transfer the rich and varied multimodal inputs they have on digital platforms to the highly structured and formal written products.

References

- Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2019). Organisational skills in academic writing: A study on coherence and cohesion in Pakistani research abstracts. *Languages*, 4(4), 1–26. <https://doi.org/10.3390/languages4040092>.
- Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2025). Determining the L2 academic writing development stage: A corpus-based research on doctoral dissertations. *International Review of Applied Linguistics in Language Teaching*, 63(1), 413–431. <https://doi.org/10.1515/iral-2023-0028>.
- Ahmed, H. N. (2025). *Pedagogy and cognition: A study of English grammar teaching practices at the undergraduate level in Pakistan* (Doctoral Dissertation). Department of English Language and Literature, The University of Lahore, Pakistan.
- Ameen, M., Saleem, N., & Taj, I. (2023). Investigating the role of English as a second language in supporting career development at tertiary level education in Pakistan. *VFAST Transactions on Education and Social Sciences*, 11(1), 180–191. <https://doi.org/10.21015/vtess.v11i1.1451>.
- Anson, C. M., Dannels, D. P., & St. Clair, K. (2005). Teaching and learning a multimodal genre in a psychology course. In A. J. Herrington & C. Moran (Eds.), *Genre across the Curriculum* (pp. 171–195). Logan: Utah State University Press.
- Cope, B., & Kalantzis, M. (2016). *E-learning ecologies: Principles for new learning and assessment*. New York: Routledge.
- Das, R., & Singh, T. D. (2023). Multimodal sentiment analysis: A survey of methods, trends, and challenges. *ACM Computing Surveys*, 55(13s), 1–38. <https://doi.org/10.1145/3586075>.
- Fareed, M., Ashraf, A., & Bilal, M. (2016). ESL learners' writing skills: Problems, factors and suggestions. *Journal of Education and Social Sciences*, 4(2), 81–92.
- Hicks, T. (2014). Adding the “digital layer”: Examining one teacher's growth as a digital writer through an NWP Summer Institute and beyond. In K. E. Pytash & R. E. Ferdig (Eds.), *Exploring Technology for Writing and Writing Instruction* (pp. 345–357). Hershey: IGI Global.
- Keane, K. (2025). Text-based discussion interactions and engagement with three different text types. *The Journal of Educational Research*, 118(2), 185–199. <https://doi.org/10.1080/00220671.2024.2423177>.
- Khan, H., Mushtaq, M., Bakhsh, K., Butt, M., & Khan, M. A. (2025). A multimodal perspective in language teaching and applied linguistics: Recent trends, issues, and future directions. *Pakistan Journal of Life & Social Sciences*, 23(2), 108–119. <https://doi.org/10.57239/PJLSS-2025-23.2.009>.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.
- Leotta, P. C., & Ahmad, M. (2025). Improving writing skills in academic English through explicit instruction and process writing. *International Journal of Linguistics*, 17(2), 80–94. <https://doi.org/10.5296/ijl.v17i2.22811>.
- Lim, F. V., & Tan-Chia, L. (2022). *Designing learning for multimodal literacy: Teaching viewing and representing*. New York: Routledge.
- Mills, K. A., Unsworth, L., & Scholes, L. (2023). *Literacy for digital futures: Mind, body, text*. New York: Routledge. <https://doi.org/10.4324/9781003137368>.
- Nazar, S., & Nordin, N. R. M. (2024). Grammaticality in writing skills of L2 English learners:

- Challenges in Pakistani academic setting. *Evolutionary Studies in Imaginative Culture*, 8(1), 517–533.
- Park, K. K. (2023). *Understanding the practices of college students who speak English as a second or foreign language during an online problem-solving task* (Doctoral Dissertation). George Mason University, Virginia, US.
- Serafini, F. (2015). Multimodal literacy: From theories to practices. *Language Arts*, 92(6), 412–423. <https://doi.org/10.58680/la201527389>.
- Siddiqui, K. A., Asad, M. M., & Rind, A. A. (2025). Learners' attitudes toward language education based on cultural identities: Global citizenship agenda in higher education institutions of Pakistan. *Journal of Applied Research in Higher Education*, 17(5), 1708–1727. <https://doi.org/10.1108/JARHE-01-2024-0034>.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471–483). New York: Routledge. <https://doi.org/10.4324/9781410612700>.
- Tan, X. (2023). Stories behind the scenes: L2 students' cognitive processes of multimodal composing and traditional writing. *Journal of Second Language Writing*, 59, 100958. <https://doi.org/10.1016/j.jslw.2022.100958>.
- Thomas, D., & Zubkov, P. (2023). Quantitative research designs. In P. Činčala, D. Penno, P. Zubkov, & S. Wa-Mbaleka (Eds.), *Qualitative Research for Practical Theology* (pp. 103–114). Berrien Springs: Andrews University and Avondale Academic Press. <https://digitalcommons.andrews.edu/books/301/>.
- Tweedale, K. (2025). Ecologies of sponsorship: Navigating the where of multiliteracies development. *Kaleidoscope: An Interdisciplinary Journal of Humanities and Social Sciences*, 1(1), 154–169. https://kaleidoscopejournal.in/current_issue.
- Unsworth, L., Cope, J., & Nicholls, L. (2019). Multimodal literacy and large-scale literacy tests: Curriculum relevance and responsibility. *The Australian Journal of Language and Literacy*, 42(2), 128-139. <https://doi.org/10.1007/BF03652032>.
- Walsh, M. (2010). Multimodal literacy: What does it mean for classroom practice? *The Australian Journal of Language and Literacy*, 33(3), 211–239. <https://doi.org/10.1007/BF03651836>.
- Yi, Y. (2014). Possibilities and challenges of multimodal literacy practices in teaching and learning English as an additional language. *Language and Linguistics Compass*, 8(4), 158–169. <https://doi.org/10.1111/lnc3.12076>.
- Zhang, E. D., & Yu, S. (2023). Implementing digital multimodal composing in L2 writing instruction: a focus on developing L2 student writers. *Innovation in Language Learning and Teaching*, 17(4), 769–777. <https://doi.org/10.1080/17501229.2022.2135712>.