

Psycholinguistic Capabilities of Students: A Measurement Construct

1 M. Farooq Tariq Butt

farooq.tariq@ucp.edu.pk

University of Central Punjab

2 Ahmed F. Siddiqi

dr.ahmedsiddiqi@ucp.edu.pk

University of Central Punjab

3 Faiza Ghulam Nabi (Corresponding Author)

faiza.tanzeem@ucp.edu.pk

University of Central Punjab

Abstract

This study develops and validates a multidimensional construct of psycholinguistic capabilities in the context of higher education. It integrates a PRISMA-guided systematic literature review of 68 high-quality studies with original empirical data from Pakistani universities. This review highlights the lack of integrative measurement frameworks that connect psycholinguistic capabilities to classroom engagement and academic success, specifically in non-Western contexts. To address this gap, survey data were collected from 350 undergraduate students (n = 200 for exploratory analyses and n = 150 for confirmatory analyses) enrolled in Business Administration and Science programs. Exploratory Factor Analysis (EFA) employing principal component analysis with Varimax rotation identified a three-factor structure i.e., Academic Expressive Pragmatic Competence, Academic Language Comprehension and Metalinguistic Awareness, and Verbal Working Memory and Integrative Processing accounting for 68.54% of the total variance. Confirmatory Factor Analysis (CFA) conducted on an independent sample affirms the strength of the construct, as evidenced by strong model fit indices, high standardized loadings, and acceptable levels of convergent and discriminant validity. The findings further indicate discipline-specific patterns e.g., students in the sciences prioritized cognitive–memory processing, whereas those in Business Administration exhibited stronger expressive–pragmatic competencies. The validated scale offers a reliable measurement tool for future structural equation modeling and broadens the scope of psycholinguistic research to include underrepresented educational contexts.

Keywords: psycholinguistic capabilities; classroom engagement; academic success; scale development; higher education

INTRODUCTION

In contemporary higher education environments, classrooms serve as complex communicative spaces in which learning primarily occurs through language-rich interactions rather than passive content delivery. In multilingual and post-colonial contexts, students are required to engage in complex academic discourse, interpret abstract explanations, participate in collaborative activities, and articulate sophisticated ideas under cognitively demanding conditions. As pedagogical strategies evolve towards discussion-based, problem-centered, and digitally mediated instruction, language has emerged as the primary medium through which learning is accessed, experienced, and enacted.

Despite this reality, academic achievement continues to be largely assessed using conventional metrics such as grades, content mastery, and general cognitive ability. While these metrics are valuable, they often obscure the language-based cognitive processes that enable students to comprehend instructions, engage in classroom discourse, and fulfill academic requirements. Consequently, challenges in engagement or comprehension are frequently attributed to low motivation, insufficient effort, or inadequate prior knowledge, while underlying psycholinguistic limitations remain unrecognized. This misalignment perpetuates a persistent gap between institutional perceptions of learning difficulties and students' actual experiences in real classroom settings. The need to explore psycholinguistic abilities is especially urgent in higher



education systems with linguistic diversity, such as those in many South Asian universities, where teaching often occurs in a second or third language. In these settings, classroom involvement includes not just active participation but also the students' ability to comprehend, assimilate, and express meaning using academic language. However, institutional strategies frequently focus on surface-level solutions, like remedial language classes or communication skills workshops, without tackling the core cognitive-linguistic processes essential for learning.

Current studies often regard language as either a technical ability or a teaching tool, which fragments its role and neglects its function as a cohesive system for cognition, communication, and engagement. As a result, psycholinguistic processes have primarily been examined in clinical settings or early education, rather than as essential skills that influence engagement and achievement in higher education (Nurjanah, N. et al., 2024). In order to assess the significance of psycholinguistic tools, a systematic review of six prominent databases was conducted that identified 43 SE-related papers employing LIWC. The findings assessed a number of applications, that include examining team interaction to identify emotional features in a personality facilitating in developing ML models to infer deleted Stack Overflow and comparing AI-generated and human-written text (Sajadi et al., 2025). Another recent research was conducted based on systematic review that identified key trends, themes and expansions in the field of psycholinguistics and treatments concerning to mental health (Overchuk et al., 2023). In order to assess human mind possibilities one of the largest coordinate-based meta-analysis of language processing was conducted that was consisted of 403 experiments (Turker et al, 2023). Hence, there is an intensified need and emergence of interdisciplinary and data-driven programs that place significant psycholinguistic demands on individuals, such as the need for continuous verbal reasoning, interpretation of technical language, and proficient academic communication. While there is an increasing awareness about classroom engagement and academic success to be fundamentally linked to language-mediated cognitive processes, empirical studies that merge psycholinguistic perspectives with classroom engagement research remain limited, specifically outside Western contexts. This gap is further widened by the absence of validated, multidimensional tools that can effectively capture psycholinguistic capabilities as experienced by students in real classroom settings.

The current research fills existing gaps by focusing on psycholinguistic capabilities as a key element in classroom learning. Utilizing psycholinguistic theory, it conceptualizes these capabilities as an underlying, multifaceted system of language-related cognitive processes that facilitate understanding, involvement, and ongoing engagement. Specifically, the study (a) defines psycholinguistic capabilities within the context of classroom learning, (b) identifies dimensions grounded in theory, (c) creates a measurement tool using a 5-point Likert scale, and (d) details a thorough process for scale development and validation through exploratory and confirmatory factor analysis. By adopting this approach, the study enhances a comprehensive understanding of classroom engagement and student achievement in modern higher education. Psycholinguistics examines the cognitive processes that reinforce language comprehension, production, and utilization (Levitt, 1989). Within educational settings, these processes influence how learners interpret instructional input, integrate linguistic information with existing knowledge, manage cognitive load, and articulate understanding through spoken or written discourse. From an educational standpoint, psycholinguistic abilities can be viewed as relatively stable yet developable learner capacities that facilitate effective language-mediated cognition in instructional contexts.

This conceptualization is informed by several complementary theoretical traditions. Information Processing Theory (Atkinson & Shiffrin, 1968; Miller, 1956) emphasizes the limitations of human cognitive systems, particularly how attentional control and memory capacity restrict language processing during learning. Working Memory Theory (Baddeley, 2003) further clarifies the roles of the phonological loop and central executive in maintaining and manipulating linguistic information amid instructional complexity. From a comprehension viewpoint, Kintsch's (1998) Construction-Integration Model clarifies how learners form coherent mental representations by integrating lexical, syntactic, and semantic information throughout extended discourse. This model holds particular significance for academic learning, where students are required to process dense explanations and link ideas over time. Complementing these cognitive perspectives, sociocultural theory (Vygotsky, 1978) perceives language as a mediational tool through which higher-order cognition, participation, and engagement are socially constructed within classroom interactions. Lastly, Levelt's (1989) psycholinguistic model of speech production clarifies the stages through which conceptual intentions are converted into linguistic output, highlighting the cognitive demands of academic expression. Collectively, these theories position psycholinguistic capabilities as integrative mechanisms that translate instructional language into meaningful learning across comprehension, memory, inference, and production processes.

Despite strong theoretical foundations and a growing empirical literature, psycholinguistic capabilities remain underexplored as a holistic construct in classroom learning research. Existing instruments are often clinically oriented, age-specific, or limited to isolated linguistic skills, restricting their applicability to higher education settings. A clear need therefore exists for a theoretically grounded, multidimensional, and psychometrically validated measurement instrument that captures psycholinguistic capabilities as antecedents of classroom engagement and student success. The present study addresses this gap by integrating psycholinguistic theory with systematic empirical synthesis to advance both conceptual clarity and methodological innovation.

SYSTEMATIC LITERATURE REVIEW METHODOLOGY (PRISMA FRAMEWORK)

To ensure theoretical rigor and methodological transparency, this study conducted a systematic literature review guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework. PRISMA was chosen for its established use in educational psychology, applied linguistics, and learning sciences to minimize selection bias, enhance replicability, and ensure comprehensive coverage.

IDENTIFICATION PHASE

Psycholinguistic capabilities refer to a learner's cognitive-linguistic abilities that facilitate the efficient comprehension, processing, interpretation, and production of language, enabling meaningful engagement and successful learning in classroom settings. A comprehensive search was conducted across five principal databases such as Scopus, Web of Science, ERIC, PsycINFO, and Google Scholar. The search strategy employed keywords that encompassed terms related to psycholinguistic capabilities, academic language processing, verbal working memory, metalinguistic awareness, classroom discourse, student engagement, and academic success. This search process resulted in the identification of 1,248 records, with an additional 37 records located through manual examination of reference lists and key review articles, culminating in a total of 1,285 records.

SCREENING PHASE

After importing records into a reference management system, 312 duplicates were removed, resulting in 973 unique records. During the title and abstract screening, 621 records were excluded

for focusing on clinical or pathological contexts, early childhood education, non-classroom settings, or constructs unrelated to language-mediated cognition. This process retained 352 articles for full-text assessment.

ELIGIBILITY AND INCLUSION

During the full-text screening, predefined inclusion criteria were applied, focusing on contexts within higher education or secondary classrooms and emphasizing theoretical or empirical relevance to psycholinguistic processes, as well as connections to learning, engagement, or performance outcomes. Ultimately, 241 articles were excluded due to limited relevance, narrow skill-based focus, insufficient methodological rigor, or weak construct operationalization. Sixty-eight studies were included in the qualitative synthesis. These studies formed the empirical and theoretical basis for conceptualizing psycholinguistic capabilities as a multidimensional construct and guided the development of the measurement instruments.

FIGURE 1. Multidimensional Construct



The synthesis identified six prominent, frequently discussed dimensions pertaining to the multidimensional construct: (i) Language Comprehension Efficiency, (ii) Working Memory for Language, (iii) Lexical and Semantic Access, (iv) Metalinguistic Awareness, (v) Language Production and Expressive Clarity, and (vi) Inferencing and Pragmatic Competence. Although these dimensions are interrelated, they constitute significant pillars of students' psycholinguistic capabilities in higher education.

INSTRUMENT DEVELOPMENT: PSYCHOLINGUISTIC CAPABILITIES QUESTIONNAIRE (PCQ)

Scale development is a systematic, theory-driven process that translates abstract constructs into reliable and valid measurable indicators. The development of the Psycholinguistic Capabilities Questionnaire (PCQ) was developed with clear construct conceptualization, including specification of conceptual boundaries and differentiation from related constructs.

SCALE VALIDATION AND STRATEGY

All items are measured on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). Drawing on the six dimensions identified in the literature synthesis, a concise 20-item questionnaire was developed. Items were adapted from previously validated instruments reported in the 68 reviewed studies to preserve conceptual validity and reliability. Particular attention was given to minimizing respondent burden while ensuring comprehensive coverage of the construct domain. The complete instrument appears in the Appendix (Table 2).

DATA COLLECTION AND ITEM ANALYSIS

Items were reviewed for clarity, academic relevance, and linguistic neutrality. Content validity was assessed through expert judgment from scholars in psycholinguistics and education. Cognitive

interviews with students further confirmed that items were interpreted as intended. The data for this study were configured through a carefully planned, two-episode collection design to ensure methodological rigor, construct validation, and robustness of the findings. Data collection was carried out in the closing days of 2025, with each episode serving a distinct analytical purpose aligned with best practices in scale development and latent variable modeling.

In the first episode, data were collected from 200 higher education students using the full set of items included in the proposed psycholinguistic capabilities questionnaire. This dataset was exclusively employed for Exploratory Factor Analysis (EFA) with the objective of uncovering the underlying dimensional structure of the construct without imposing any a priori constraints. The sample size was considered adequate for exploratory purposes and allowed for reliable estimation of inter-item correlations and communalities. The respondents in this phase comprised of an equal number of male and female students, drawn from Science and Business Administration programs across multiple private-sector universities located in Lahore. This heterogeneity was intentionally established to enhance diversity and improve the generalizability of the extracted factor structure across analytically demanding academic contexts.

In the second episode, a new and independent sample of 150 students was collected from a similar population i.e., students enrolled in Science and Business Administration programs at private universities in Lahore. This dataset was reserved exclusively for Confirmatory Factor Analysis (CFA) to validate the factor structure identified during the exploratory phase. By using a different, yet demographically and academically comparable sample, the study minimized the risk of capitalization on chance and strengthened the credibility of the measurement model. For example, in the first episode, it was made sure that gender representation should remain balanced, ensuring equal participation of male and female students, thereby reducing gender-based sampling bias in the validation process.

Several procedural controls were embedded in the data configuration to enhance data quality. Respondents were engaged individually and independently, with particular care taken to avoid group-based or classroom-administered responses that could artificially inflate inter-item correlations. This approach helped mitigate potential multicollinearity arising from shared response contexts. Furthermore, the entire data collection for each episode was completed within a single week, which reduced the likelihood of auto-correlation effects stemming from temporal shifts in academic workload, assessment pressure, or instructional variation.

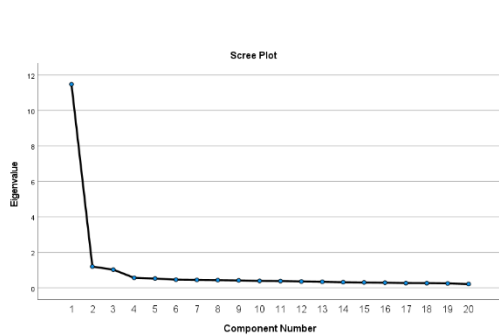
Overall, this two-stage data configuration that combines exploratory and confirmatory samples, balanced gender representation, disciplinary diversity, and procedural controls, provides a strong empirical foundation for the study. It aligns with established methodological standards for scale development and validation, ensuring that the reported psycholinguistic capability dimensions are both empirically sound and generalizable within the context of private-sector higher education institutions in Lahore.

THREE FACTOR ANALYSIS

In order to explore psycholinguistic possibilities, a comprehensive analysis was conducted that comprised of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). The three-factor analysis via Exploratory Factor Analysis (EFA) indicated the actual proportion of variance and represents expressive-pragmatic academic language competence whereas the Confirmatory Factor Analysis (CFA) was conducted for reliability and the validation of the construct that was developed earlier.

EXPLORATORY FACTOR ANALYSIS (EFA)

EFA was conducted on the 20 items using principal component analysis with Varimax rotation. Prior to analysis, items were scrutinized for distributional properties and inter-item correlations. No items exhibited extreme skewness or kurtosis, and all inter-item correlations were adequate. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.971, indicating excellent suitability for factor analysis. Bartlett's Test of Sphericity was highly significant ($\chi^2(190) = 4319.506, p < .001$), confirming that the correlation matrix differed significantly from an identity matrix.



KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.971
Bartlett's Test of Sphericity	Approx. Chi-Square	4319.506
	Df	190
	Sig.	.000

FIGURE 2. KMO and Bartlett's Test

The analysis yielded a clear three-factor solution with no problematic cross-loadings and strong communalities (most > 0.60). The factors accounted for 68.54% of the total variance (Factor 1: 30.51%, Factor 2: 23.66%, Factor 3: 14.37%). Table (4.1) presents the rotated component matrix.

TABLE 1. Rotated Component Matrix (Principal Component Analysis, Varimax Rotation)

Community		1	2	3
	Percentage of Variance	30.51	23.66	14.37
	Cummulative Percentage	30.51	54.17	68.54
0.713	I can easily find the right words to express my academic ideas.	.775		
0.687	I can infer what my instructor expects even when it is not stated directly.	.728		
0.675	I understand subject-specific terminology without much difficulty.	.724		
0.639	I can clearly explain my ideas during class discussions.	.722		
0.665	I understand implied meanings in classroom discussions.	.699		
0.646	I can adapt my responses based on the reactions of classmates and instructors.	.696		
0.662	I can quickly grasp the meaning of new academic words from context.	.692		
0.661	I can interpret questions and feedback in a constructive way.	.670		
0.613	I can express complex ideas in simple and understandable language.	.666		
0.661	I can organize my thoughts effectively when speaking in class.	.652		
0.687	I can quickly understand the main ideas when my instructor explains a concept verbally.		.765	
0.678	I can identify when language is ambiguous or unclear in lectures or texts.		.734	
0.717	I can follow long explanations without needing repeated clarification.		.726	
0.722	I easily understand academic texts used in my courses.		.721	

0.686	I rarely feel lost when listening to complex classroom discussions.	.693	
0.673	I consciously adjust my language depending on the academic context.	.648	
0.667	I am aware of how sentence structure affects meaning.	.565	
0.795	I can remember key points from lectures while listening to new information.		.822
0.743	I can hold multiple ideas in mind while solving language-based problems.		.781
0.716	I can connect earlier parts of a lecture to later explanations.		.701

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

The first factor explains the largest proportion of variance (approximately 30.51%) and represents expressive–pragmatic academic language competence. Items loading on this factor reflect students’ ability to access vocabulary effortlessly. It further reflects how well students can organize ideas in a coherent manner, and adapt language use in actual classroom environment. This dimension not only represents linguistic production but also pragmatic employability of a language in real time interactions. It further includes awareness of audience reactions and interpretation of implied meanings. Conceptually, it reflects how actively students employ language in order to participate in various speech events and discussions with an ability to negotiate meaning, and construct knowledge in social learning environments. Grounded in pragmatics theory, communicative competence, and sociocultural perspectives, this factor highlights language as a tool for interaction rather than mere transmission of information. High scores on this dimension indicate students who are verbally more engaged and can articulate complex ideas with clarity. Furthermore, they are more responsive to the instructional cues, making this factor as a strong predictor of observable classroom participation and interactive learning.

The second factor, that accounts for approximately 23.66% of the variance, signifies academic language comprehension and metalinguistic awareness. This dimension includes students’ capacity to comprehend spoken and written academic language, identify and decode complex ideas, and consciously reflect on how language structure influences meaning. Unlike the first factor, which focuses upon outward expression, this factor reflects internal, cognitive, and reflective processes involved in decoding and observing linguistic input. It aligns closely with psycholinguistic processing models, theories of metalinguistic awareness, and cognitive load theory, emphasizing efficient comprehension under highly cognitive demanding conditions. Students who score high on this factor are able to understand extended explanations, comprehend complex academic texts, and identify ambiguities early, ultimately reducing the need for repeated clarification. This dimension is foundational for conceptual understanding and plays a critical role in academic success, specifically in lecture-driven and text-intensive learning environments.

The third factor explicates approximately 14.37% of the variance and captures verbal working memory and integrative processing. Items loading on this factor strongly reflect students’ ability to store multiple pieces of linguistic information in mind, integrate initial and later parts of a particular lesson, and organize sequential information processing. Overall, this factor represents the cognitive infrastructure that supports both comprehension and expression by enabling online integration of information over time. Theoretical grounding in working memory models, capacity theories of comprehension, and information processing frameworks highlight its role in sustaining understanding during complex and cumulative academic tasks. Students with high scores on this dimension are better equipped to cope with complex lessons, connect concepts across time, and perform effectively in highly demanding problem-solving contexts involving high cognitive

processing which makes this factor specifically noticeable for disciplines such as science and technology.

Taken together, the three-factor structure exposes psycholinguistic capacity in academic contexts not a unitary skill but a multi-layered and interdependent mechanism. Verbal working memory and integrative processing provide the cognitive capacity essential for efficient learning, while academic language comprehension and tasks related to metalinguistic awareness enable students to accurately decode and observe linguistic input. These two dimensions, in turn, support expressive-pragmatic competence, which translates internal comprehension into significant classroom engagement and learning. This hierarchical relationship, aligning closely with contemporary perspectives, view learning as both a cognitive and communicative process. Overall, the findings demonstrate that students’ academic language performance emerges from the interaction of cognitive capacity, receptive processing, and expressive-pragmatic use, offering a vigorous and theoretically meaningful framework for understanding psycholinguistic functioning in higher education classrooms.

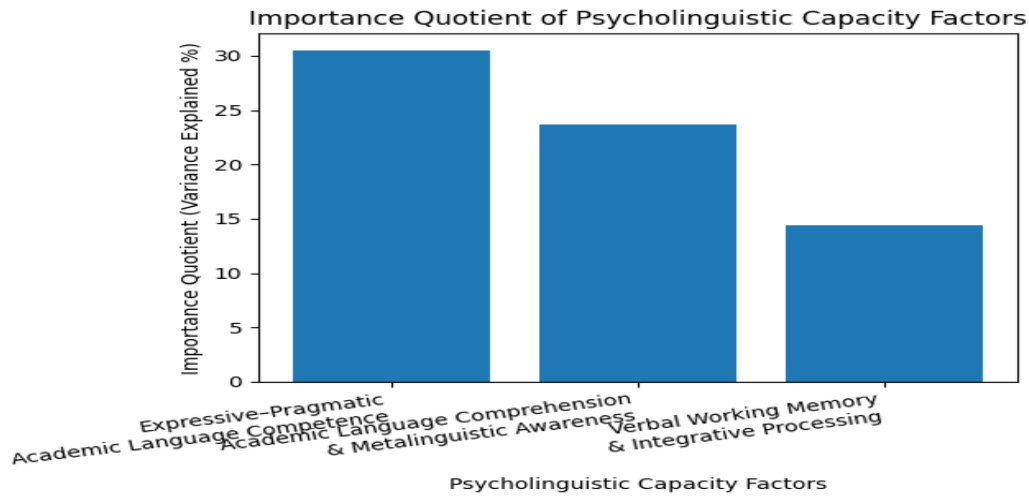


FIGURE 3. Visual Representation of Psycholinguistic Capacity Factors
CONFIRMATORY FACTOR ANALYSIS (CFA)

The structure explored in the previous section needed to be confirmed; preferably with a different sample; however, with identical characteristics. Thus, the confirmation is made both for reliability and the validation of the construct developed in the previous sections. CFA should be performed on an independent sample to validate the factor structure. Model fit may be assessed using: - CFI and TLI ≥ 0.90 - RMSEA ≤ 0.08 - SRMR ≤ 0.08 and a few other ratios, as defined and discussed by Byrne, B. M. (2001).

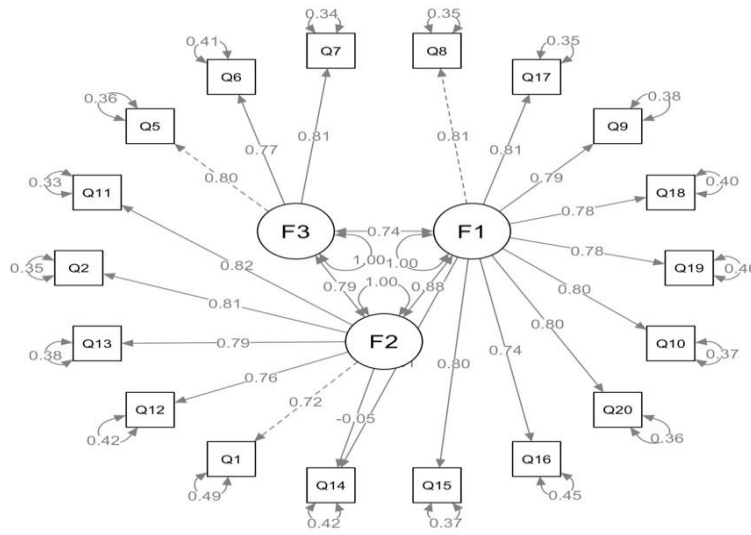


FIGURE 4. Path Diagram for CFA
RESULTS AND DISCUSSION

Different ratios and indices to highlight and ensure the applicability of the CFA are represented here. $\chi^2/df = 1.4111$, RMSEA = 0.048 (with 90% CI: {0.034, 0.062}), PCLOSE = 0.562, GFI = 0.933, AGFI = 0.906, PGFI = 0.672, CFI = 0.962, NFI = 0.907, RFI = 0.886, IFI = 0.962, TLI = 0.953, PRATIO = 0.817, PNFI = 0.740, PCFI = 0.785, NCP = 60.511 (with 90% CI: {29.983, 98.953}), FMIN = 0.6, F-0 = 0.229 (with 90% CI: {0.114, 0.375}), ECVI = 0.888 (with 90% CI: {0.773, 1.034}), MECVI = 0.908, AIC = 234.511 (Saturated Model: 272.000, Independent Model: 1728.728), BCC = 239.742 (Saturated Model: 290.721, Independent Model: 1730.931), BIC = 475.900 (Saturated Model: 1,135.915, Independent Model: 1830.365), CAIC = 408.541 (Saturated Model: 894.843, Independent Model: 1802.004)

Convergent validity (AVE, CR) and discriminant validity (Fornell–Larcker criterion) should be established, factor wise.

	F1	F2	F3
alpha	0.9410623	0.8888920	0.8372292
omega	0.9419849	0.8597669	0.8368237
omega2	0.9525698	0.6345683	0.8368237
omega3	0.9526550	0.6344566	0.8361183
avevar	NA	NA	0.6310069

The numerals reported for Factors 1–3 represent direct and substantive evidence for convergent validity, as well as strong support for the measurement quality of the psycholinguistic capacity scale that was developed in this investigation. Convergent validity refers to the extent to which items intended to measure the same latent construct are strongly interrelated and jointly represent that construct. In this context, the reported reliability and variance-based indices demonstrate that the items within each factor converge meaningfully on their intended psycholinguistic dimension. The Cronbach’s alpha coefficients for the three factors are uniformly high (F1 = 0.94, F2 = 0.89, F3 = 0.84), showcasing strong internal consistency among the items within each factor. From a convergent validity perspective, these values suggest that the items designed to measure each



psycholinguistic dimension share substantial common variance and operate cohesively rather than independently. For this investigation, such high alpha values indicate that expressive–pragmatic competence, academic language comprehension, and verbal working memory are each measured by internally coherent item sets, reflecting successful operationalization of the theoretical constructs.

The McDonald’s omega coefficients further strengthen this conclusion. Omega values for all three factors ($F1 \approx 0.94$, $F2 \approx 0.86$, $F3 \approx 0.84$) assure that the observed reliability is not only a function of item redundancy but it also reflects true-score variance attributable to the latent constructs for omega is less sensitive to violations of tau-equivalence and therefore, its consistency with alpha provides clear evidence that the indicators within each factor genuinely converge on a common underlying psycholinguistic process. In the context of this study, omega supports the claim that each factor represents a constant and meaningful component of psycholinguistic capacity.

The omega hierarchical variants (omega2 and omega3) offer deeper insight into the strength and purity of convergence. For Factor 1, very high omega2 and omega3 values (both ≈ 0.95) underscore the variance in item responses overwhelmingly explained by a single dominant latent dimension. This specifically provides strong convergent validity evidence for expressive–pragmatic academic language competence, suggesting that its items converge tightly on a well-defined construct. For Factor 2, lower omega2 and omega3 values (≈ 0.63) suggest that while the items converge sufficiently to form a reliable factor, the construct is cognitively more complex and may involve overlapping sub-processes, such as comprehension, ambiguity detection, and metalinguistic reflection. This pattern is theoretically consistent with the multifaceted nature of academic language comprehension. Factor 3 again shows strong convergence (≈ 0.84), indicating that verbal working memory and integrative processing are measured with substantial coherence. The Average Variance Extracted (AVE) value, reported for Factor 3 (≈ 0.63), exceeds the commonly accepted threshold of 0.50, indicating that more than half of the variance in the indicators is captured by the latent construct rather than by measurement error. This constitutes direct statistical confirmation of convergent validity for this factor. Although AVE is not reported for Factors 1 and 2 in the table, the combination of very high factor loadings, strong communalities, and high omega coefficients strongly implies that AVE would also meet or exceed acceptable levels for these dimensions.

Taken together, these numerals validate that the psycholinguistic capacity scale exhibits strong convergent validity across all three factors. The items within each dimension cohere statistically and conceptually, reflecting successful translation of theoretical psycholinguistic constructs into empirical indicators. For the present investigation, this means that expressive–pragmatic language use, academic language comprehension, and verbal working memory are not only theoretically distinct but are also measured with precision, consistency, and substantive convergence, providing a sound foundation for subsequent structural modeling and substantive interpretation of students’ psycholinguistic capabilities.

> discriminantValidity(fit)

lhs	op	rhs	est	ci.lower	ci.upper	Df	AIC	BIC	Chisq	Chisq diff	RMSEA	Df diff
Pr(>Chisq)												
1	F1 ~	F2	0.8789926	0.8407448	0.9172403	132	6701.842	6846.290	156.0478	1.308006		
			0.03204197	1	2.527559e-01							
2	F1 ~	F3	0.7434276	0.6771339	0.8097214	132	6739.300	6883.747	193.5051	38.765388		
			0.35480224	1	4.779277e-10							

3 F2 ~ F3 0.7887172 0.7255104 0.8519241 132 6718.565 6863.013 172.7703 18.030569
0.23826155 1 2.173861e-05

The numerals reported in the discriminant validity output provide clear, demanding, and statistically defensible evidence regarding the distinctiveness of the three psycholinguistic capacity factors examined in this investigation. Discriminant validity refers to the extent to which theoretically distinct constructs are empirically separable, even when they are conceptually related. In the context of psycholinguistic capacity, this is especially important because expressive competence, comprehension, and working memory are expected to correlate, but not collapse into a single undifferentiated ability.

The estimated latent correlations between the factors are all positive and statistically meaningful, which is theoretically appropriate given that the three dimensions jointly form a higher-order psycholinguistic system. The correlation between Factor 1 and Factor 2 (0.879) indicates a strong relationship between expressive–pragmatic academic language competence and academic language comprehension and metalinguistic awareness. However, the confidence interval for this estimate (0.841 to 0.917) does not include unity (1.0), suggesting that while these constructs are closely related, they are not empirically identical. The associated chi-square difference test is non-significant ($\Delta\chi^2 = 1.308$, $p = .253$), signifying that constraining these two factors to be flawlessly correlated does not significantly improve model fit. In the context of this investigation, this result suggests that Factor 1 and Factor 2 are highly interdependent yet theoretically and empirically distinguishable, reflecting adjacent but non-redundant layers of psycholinguistic functioning.

The relationship between Factor 1 and Factor 3 (0.743) provides stronger evidence of discriminant validity. The confidence interval (0.677 to 0.810) is comfortably below 1.0, and the chi-square difference is large and highly significant ($\Delta\chi^2 = 38.77$, $p < .001$). This indicates that forcing these two factors to be identical would significantly worsen model fit. For the present investigation, this demonstrates that expressive–pragmatic competence and verbal working memory represent clearly distinct components of psycholinguistic capacity, despite being functionally related during classroom learning. Similarly, the correlation between Factor 2 and Factor 3 (0.789) shows a substantial but bounded association. The confidence interval (0.726 to 0.852) again excludes unity, and the chi-square difference test is significant ($\Delta\chi^2 = 18.03$, $p < .001$). This result confirms that academic language comprehension and verbal working memory, while cognitively intertwined, are empirically separable constructs. In the context of this study, this supports the view that comprehension processes rely on, but are not reducible to, working memory capacity.

Taken together, these results provide strong support for discriminant validity of the three-factor psycholinguistic capacity model. The latent correlations demonstrate theoretically expected relatedness, while the confidence intervals and chi-square difference tests confirm that none of the factor pairs are statistically indistinguishable. For this investigation, this pattern is particularly meaningful because it validates the proposed layered structure of psycholinguistic capacity: working memory supports comprehension, comprehension enables expressive–pragmatic competence, yet each dimension retains its own empirical identity. Consequently, the scale does not merely measure a general language ability but captures distinct, theoretically grounded components of psycholinguistic functioning in academic contexts, strengthening the overall construct validity of the measurement model.

CONCLUSION

This study was motivated by a persistent gap in the higher education research. For example, student engagement and success are often explained through motivation, intelligence, or content mastery, while the psycholinguistic mechanisms that mediate instruction remain largely unexamined. In linguistically complex environments, such as English-medium instruction in Pakistan, misalignment between instructional demands and students' psycholinguistic capabilities can lead to disengagement or underperformance.

Guided by psycholinguistic processing theory, communicative competence, sociocultural learning theory, and cognitive capacity models, this study employed a rigorous two-stage scale development and validation process. The EFA uncovered a robust three-factor structure that integrates cognitive, receptive, and expressive processes. CFA on an independent sample confirmed the model's stability, strong fit and validity. These findings extend classical theories (e.g., Baddeley, Levelt, Vygotsky) into contemporary classroom engagement research and challenge models that treat language as a neutral medium for communication.

The validated PCQ offers a practical, multidimensional tool for educational researchers and practitioners. Discipline-specific patterns further support domain-specific cognition theory: science programs place heavier demands on working memory and comprehension, whereas business programs emphasize expressive-pragmatic skills. Limitations include the cross-sectional design, reliance on self-report measures, and the sample's restriction to private universities in Lahore, Pakistan. Future research should employ longitudinal designs, experimental interventions, cross-national validation, and integration with learning analytics or AI platforms.

The findings are particularly relevant for higher education in Pakistan and similar multilingual contexts. By making psycholinguistic demands visible and measurable, this study equips institutions, faculty, and policymakers to design cognitively and linguistically inclusive curricula, assessments, and teaching strategies. Globally, it highlights that student success depends not only on what is taught, but also on how language and cognition jointly shape access to learning. Ultimately, the PCQ positions psycholinguistic capabilities as a critical, yet previously missing, link between teaching, engagement, and academic success.

The implications extend beyond classroom scenarios to academic policies and educational strategies across the nation. In multilingual societies, where linguistic diversity often intersects with socioeconomic disparities, the PCQ offers a framework for identifying hidden barriers to learning and promoting equity. It encourages universities to rethink assessment design, ensuring that measures of knowledge are not confounded by the language processing demands. For faculty, it highlights the importance of scaffolding instruction in ways that respect cognitive and linguistic accessibility. For policymakers, this study provides evidence to support reforms that prioritize inclusive pedagogy as a cornerstone of quality education.

At the global level, this study contributes to the discourse on educational justice by demonstrating that psycholinguistic competence is not peripheral but central to academic achievement. It reframes student performance as a function of both intellectual engagement and linguistic mediation, urging educators to recognize language as a cognitive tool, rather than a neutral medium. This bridges the gap between psycholinguistics and pedagogy, offering a replicable model for diverse educational systems. In conclusion, the PCQ specifically excels in causing a paradigm shift. The success rate in higher education is not only about transmitting knowledge but also about cultivating environments where language and cognition are harmonized to unlock

learning potential. This insight challenges institutions worldwide to move toward more inclusive, evidence-based practices that honor the complexity of students learning.

References

- Nurjanah, N., Mizan, A., Tamami, M. Z., & Supriatna, N. K. (2024). Understanding Child language development patterns based on developmental psychology and psycholinguistics approaches. *International Journal of Language and Culture*, 2(2), 68-77.
- Sajadi, A., Damevski, K., & Chatterjee, P. (2025). Psycholinguistic Analyses in Software Engineering Text: A Systematic Literature Review. *arXiv preprint arXiv:2503.05992*.
- Overchuk, V., Vdovichenko, O., Kovalova, O., Harbera, I., & Pylypak, V. (2023). The role of psycholinguistics in understanding the historical development of methods of treatment of mental disorders. *Research Journal in Advanced Humanities*, 4(1), 103-117.
- Turker, S., Kuhnke, P., Eickhoff, S. B., Caspers, S., & Hartwigsen, G. (2023). Cortical, subcortical, and cerebellar contributions to language processing: A meta-analytic review of 403 neuroimaging experiments. *Psychological Bulletin*, 149(11-12), 699.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839. <https://doi.org/10.1038/nrn1201>
- Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511605963>
- Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International journal of testing*, 1(1), 55-86.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387. <https://doi.org/10.2307/356600>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- Gombert, J. E. (1992). *Metalinguistic development*. University of Chicago Press.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). Academic Press.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149. <https://doi.org/10.1037/0033-295X.99.1.122>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511805875>
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97. <https://doi.org/10.1037/h0043158>
- Nagy, W. E. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 64-83). Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4
- Tunmer, W. E., & Herriman, M. L. (1984). The development of metalinguistic awareness: A conceptual overview. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic awareness in children* (pp. 12-35). Springer. https://doi.org/10.1007/978-3-642-69113-3_2
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Questionnaire (Appendix)

Table 2 Psycholinguistic Capabilities Questionnaire (PCQ)

	Mean	Skew	Kurt
Dimension 1: Language Comprehension Efficiency			
1. I can quickly understand the main ideas when my instructor explains a concept verbally. (Just & Carpenter, 1992)	2.79	-0.01	-0.24
2. I rarely feel lost when listening to complex classroom discussions. (Kintsch, 1998)	2.76	-0.09	-0.14
3. I can follow long explanations without needing repeated clarification. (Baddeley, 2003)	2.85	-0.11	-0.14
4. I easily understand academic texts used in my courses. (Nation, 2001)	2.78	-0.27	0.22
Dimension 2: Working Memory for Language			
5. I can remember key points from lectures while listening to new information. (Baddeley, 2003)	2.80	0.15	0.00
6. I can hold multiple ideas in mind while solving language-based problems. (Miller, 1956)	2.88	0.01	0.74
7. I can connect earlier parts of a lecture to later explanations. (Kintsch, 1998)	2.84	0.29	0.20
Dimension 3: Lexical and Semantic Access			
8. I can easily find the right words to express my academic ideas. (Levelt, 1989)	2.84	0.07	-0.44
9. I understand subject-specific terminology without much difficulty. (Nation, 2001)	2.83	-0.26	0.13
10. I can quickly grasp the meaning of new academic words from context. (Nagy, 1997)	2.80	0.07	-0.43
Dimension 4: Metalinguistic Awareness			
11. I am aware of how sentence structure affects meaning. (Gombert, 1992)	2.82	0.00	0.21
12. I can identify when language is ambiguous or unclear in lectures or texts. (Tunmer & Herriman, 1984)	2.80	-0.08	-0.10
13. I consciously adjust my language depending on the academic context. (Bialystok, 2001)	2.80	-0.09	-0.07
Dimension 5: Language Production and Expressive Clarity			
14. I can clearly explain my ideas during class discussions. (Levelt, 1989)	2.79	0.01	-0.28
15. I can organize my thoughts effectively when speaking in class. (Flower & Hayes, 1981)	2.80	0.07	-0.45
16. I can express complex ideas in simple and understandable language. (Sweller, 1988)	2.86	0.14	0.11
Dimension 6: Inferencing and Pragmatic Competence			
17. I can infer what my instructor expects even when it is not stated directly. (Grice, 1975)	2.80	-0.01	-0.24
18. I understand implied meanings in classroom discussions. (Sperber & Wilson, 1995)	2.79	0.18	-0.60
19. I can adapt my responses based on the reactions of classmates and instructors. (Vygotsky, 1978)	2.84	-0.02	-0.25
20. I can interpret questions and feedback in a constructive way. (Hattie & Timperley, 2007)	2.81	0.00	-0.26