© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1712260

Effectiveness Of an Immersive Flow State Approach in Reducing Speaking Latency and Improving L2 Oral Fluency

J. RAJKAMAL PETRO[1], DR. M. BEULAH HEMALATHA[2]
[1]*Ph.D. Research Scholar, PG & Research Department of English, Nazareth Margoschis College at Pillaiyanmanai, India Affiliated to Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India*
[2]*Research Guide & Supervisor, Assistant Professor of English, PG & Research Department of English, Nazareth Margoschis College at Pillaiyanmanai, India Affiliated to Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India*

Abstract- ***Indian higher education Regional-medium students are usually able to show a good level of English literacy but constant deficiency in spoken English, which is often explained by the so-called affective mind blocks anxiousness, shyness, fear of negative judgment. The research question is whether or not a pedagogical approach to modeling informed by the Flow Theory, called the Immersive Flow State Approach (IFSA), can reduce the latency in speaking and enhance the L2 oral performance in the Tamil-medium engineering undergraduates. A randomized pretest- post-test two-group experimental design was used having N = 200 Tamil-medium undergraduate engineering students in a private engineering college in Tamil Nadu, India (experimental n = 150; control n = 50). The course was 12-weeks long, combined with the principles of the Flow Theory, flipped learning, and communicative and task-based classroom practice. The speaking performance was evaluated based on the Dialogue Completion, Picture-Cued, and Responsive Assessment activities graded on the basis of a 5-point analytic rubric that included the following points: response time, fluency, relevancy/accuracy, vocabulary, and grammar; the inter-rater reliability was determined (Cohen kappa). Data analysis was done in SPSS version 28.0 with pre-post analysis and additional analyses. The procedures of ethical approval and written informed consent were observed. There were significant pre- and post-experimental gains on the results of the experimental group, such as response time/latency and fluency. As an illustration, in the Dialogue Completion Task, response-time ratings rose by M = 1.40 (SD = 0.50) between pretest and post-test and the gain in fluency by M = 1.32 (SD = 0.476) to M = 2.96 (SD = 0.611) (statistically significant, p =.001). Conversely, the control group did not demonstrate a significantly high or varied measure change. The outcomes of the learner feedback showed that the perceived changes of the anxiety-related barriers were strong and the perceptions of the elements of the intervention were positive.***

Keywords: Flow Theory, L2 Speaking Fluency And Speaking Anxiety, Flipped Classroom, Regional-Medium Learners.

I. INTRODUCTION

High-stakes oral communication in English is a quality-of-life skill among engineering undergraduates in India, which defines academic engagement, labour marketability, and access to worldwide work populations. But the most delicate skill even in the case of functional reading and writing is among many regional-medium learners, students whose previous schooling had been conducted mostly in a regional language, which is English speaking. Within these environments, students might have declarative learning of grammar and vocabulary but fail to transfer this learning into real-time output in a fluent, timely and contextually relevant manner. Another common argument in applied linguistics is that speaking is the most vulnerable to the affective constraint (e.g., anxiety, fear of negative evaluation, shyness) as it is public, fast, and hard to revise after utterance. This is to say that L2 speaking is not only a language act, but also an emotional one and a body of research has proved this to be a reliable occurrence in weaker language performance in any given situation. There are meta-analytic data that suggest anxiety has a significant relationship with achievement and performance, which makes it clear that affect is not marginal but structurally related to language development and assessment performance (Teimouri,

Goetze, and Plonsky, 2019). Notably, anxiety is both a personality characteristic and a contextual response to speaking activities, interpersonal communication, or evaluative contexts, and it therefore makes speaking performance highly susceptible to the mind block phenomenon that can be either delayed responses, lack of fluency, avoidance, or talking by memorizing bits instead of engaging in dialogue.

Theoretical explanations of classic accounts can be used to understand why anxiety interferes with the process of L2 speaking. Affectively filter perspective of viewing it would be that when anxiety is high, the learners will be less willing to take risks and can restrict their input or feedback uptake (Krashen, 1982). Socio-culturally speaking, the development of speaking requires facilitated participation and mediated interaction within the zone of proximal development of the learner (Vygotsky, 1978); but nervous learners tend to evade exactly the type of interactional opportunity that facilitates development. The fieldwork on measurement has also been keen to note that language anxiety is complex construct associated with communication apprehension and fear of negative judgment, which is particularly acute in speaking (Horwitz, Horwitz, and Cope, 1986). Although these basic frameworks are still influential, recent studies have also been led to posit that speaking development is most well understood within dynamic affective systems that encompass both debilitating emotions (e.g., anxiety) and enabling emotions (e.g., enjoyment), and that pedagogy does not just help to reduce anxiety, but that positive engagement in the process of speaking is actively developed. In this respect positive psychology in SLA has assisted in the re-framing of the classroom emotion as a re-development rather than an obstacle. Empirical studies are often finding that enjoyment and anxiety can be comorbid, however, can be linked to engagement and performance in different ways across the classroom ecology and task design. To take but an example, the research published in the top journals of applied linguistics revealed that achievement-relevant emotions: enjoyment, anxiety, and boredom can coalesce to influence the language learning experience and outcome, which means that the successful pedagogy must be aimed at levelling the emotional climate of speaking teaching instead of focusing on anxiety as an independent variable. Another closely related construct, flow, provides a rather practical perspective on instructional design of speaking to learners with inhibition issues. The concept of flow theory equates optimal experience to a state of significant involvement whereby there is a high degree of concentration, a sense of control and internal pleasure when perceived challenges are matched with perceived skills (Csikszentmihalyi, 1990). Flow in L2 classrooms has been reviewed as an experience facilitated by a task that may assist in maintaining attention, reducing self-consciousness, and enhancing the readiness to communicate. The current developments in leading journals have started to define the characteristics of tasks that are linked to flow in classrooms, which reinforced the claim that flow is not a personality phenomenon alone but can be pedagogically constructed by paying close attention to the task conditions and interaction design.

Simultaneously, the scope and the rigor of interventions targeted at the alleviation of foreign language anxiety have increased. According to a recent meta-analysis and systematic review of interventions on foreign language anxiety, pedagogical and psychological strategies can lead to significant changes in feelings of anxiety but also show the presence of inconsistent effects of interventions varying in accordance with the type of intervention, their duration, and limitations imposed by the context (Xiong & Zhang, 2024). This increasingly substantial body of evidence is essential in Indian higher-education settings where a student can be a regional-medium learner experiencing an additional cumulative load: a lack of experience in spontaneous engagement with the English language, high levels of perceived stakes, and social comparative competition in competitive tertiary education programs. It is in this wider framework that interaction-based communicative practice combined with systematic emotional support might hold particular promise in the form of pedagogical model. Communicative Language Teaching and task-based models focus on meaningful language use, negotiating meaning and repetitions, which come with better spoken fluency and confidence in the long term. Nevertheless, in the case of the learner who has learned to fear making a mistake or being viewed negatively, more of the speaking practice is not necessarily therapeutic, the greater the exposure to it

without the sense of emotional security, the more the avoidance may increase. Therefore, one of the current research directions has been design that can (a) increase high-quality speaking opportunities as well as (b) reduce affective load by means of supportive preparation, scaffolded performance conditions and autonomy-supportive learning environments.

Flipped learning has become one of the popular directions to be pursued by contemporary research. In the case of low-proficiency speakers or those who are inhibited, performance pressure can be minimized by this reallocation since privately preparing and rehearsing before speaking can reduce performance pressure. Empirical studies on the topic of Language Teaching Research indicate that flipped designs may enhance the learner engagement and educational outcomes especially when combined with in-class activities that promote participation as opposed to merely implementing technology in the conventional instructions (Chang, 2023). Though not a speaking-anxiety intervention per se, flipped learning can also act as an anxiety-buffering architecture as it enhances predictability, preparedness, and perceived control, the main components of low anxiety and high flow. The current research fits in between these areas of interest, L2 speaking anxiety, positive emotion, and task design, in a regional-medium engineering setting, Tamil Nadu, India. Based directly on the Flow Theory, Affective filter logic proposed by Krashen, and socio-cultural scaffolding, the study considers an instructional model, named Immersive Flow State Approach (IFSA). The rationale behind creating IFSA was to deal with the self-reported mind blocks of the learners (e.g., anxiety, shyness, fear) by building a learning sequence that promotes preparedness (e.g., pre-class input), psychological safety, and repetition of communicatively meaningful practice in the classroom. The intervention combines the features, which are typically related to communicative and technology-mediated pedagogy (e.g., flipped learning and interactive tasks) but puts them in the context of a logic of flow: it establishes the situation in which the challenge-skill ratio, concentration of mind, and a sense of control are more likely to arise during speaking tasks. Therefore, this paper addresses a pragmatic and research-impractical gap: although the state of anxiety-performance relationships is well known at the general level (Teimouri et al., 2019), not many studies test the theory-driven, classroom-applicable models of interventions that explicitly aim to reduce learners to the inhibited state and then to the optimal engagement to speak, especially in the case of Indian regional-medium undergraduates. Furthermore, since the speaking outcomes are multidimensional, the research not only evaluates the overall impressions of improvement but also the specific elements of performance, including response latency, fluency, and other aspects of the spoken task performance, as well as the perceptions of the learners about the decreased anxiety level and confidence increment. By doing that, it offers a piece of evidence on whether flow-based pedagogy can play the role of both affective intervention and performance-enhancement tool in a high-demand tertiary ELT environment.

II. METHODS

2.1 Participants

The participants were 200 undergraduate engineering students who were recruited in a private engineering college in Tamil Nadu, India. The purposive sampling plan was adopted to identify learners that were contextual in the study (mainly, the students attending schooling in Tamil language and mentioning obstacles to English speaking such as anxiety, shyness, fear of errors, and lack of previous experience of being in sustained English-speaking settings). Students who had any prior preparation to formal speaking training or which had high exposure to English-speaking settings were eliminated to ensure relative homogeneity in prior speaking experience. The age reported was 18-21 years. The sample was split into 99 males (66%), and 51 females (34%), and included all year of study (1st year: 20%; 2nd year: 32%; 3rd year: 28%; 4th year: 20%). With regard to medium of instruction, 90% (n = 135) students said they use Tamil as the medium of instruction and 10 percent (n = 15) English medium. After the recruitment, students were assigned to experimental group (n = 150) and a control group (n = 50). Post-selection randomization .

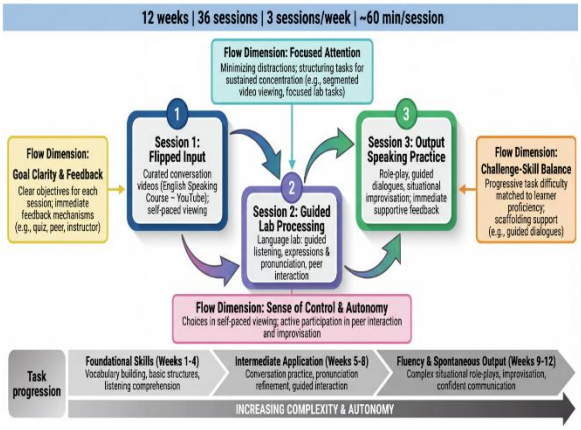
2.2 Design

The pretest- post-test design used was two-group design in which the Immersive Flow State Approach (IFSA) (experimental group) was compared with the conventional method of instruction (control group).

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1712260

The participants underwent speaking tests at baseline (pretest) and post-12 weeks of the instructional intervention (post-test). The design was aimed at the test of whether an instructional sequence with flow-theory-driven design could decrease observable hesitation (operationalized as response latency) and enhance several dimensions of performance in speech in comparison to the control condition. Since the factor of affect is highly involved in L2 performance, self-report measures (needs analysis and post-intervention feedback) were also included by the design to address the perceived barriers and learner experience. Such attention is consistent with current evidence that the language anxiety is significantly related to the performance and achievement results, which why affect-sensitive measurement should be included in the intervention study (Teimouri et al., 2019). Figure 1 demonstrates the intervention of Immersive Flow State Approach (IFSA).

Figure 1. Immersive Flow State Approach (IFSA): 12-week intervention structure and flow dimensions



The development of IFSA happened as a context-driven pedagogical framework due to the Csikszentmihalyi Flow Theory and was implemented into a designed fusion of flipped learning and communicative output-oriented classroom teaching. The intervention was specific in that it sought to minimize mind blocks (i.e. fear, shyness, anxiety) by enhancing perceived control, setting goals and matching the difficulty of the task to the emerging abilities of the learners, which has been theorized to facilitate flow-like engagement when performing learning tasks. The program was 12 weeks long and consisted of 36 learning activities, which were presented in three 60-minute sessions every week. There were two sessions in the week, which were oriented majorly towards input-oriented work, and one session in the week, where output-oriented work was focused on speaking practice. The initial weekly session was flipped in nature: the students were provided with video on conversations (a selected playlist, related to the channel "English Speaking Course" on YouTube) to watch in advance before the lesson. The second weekly meeting was in the language laboratory and was based on guided listening and peer-centered interaction oriented at expressions and pronunciation characteristics provided by the videos. The third weekly lesson was based on speaking production, with the use of structured role plays, guided dialogues and situational improvisation activities, which were based on the weekly input.

The types of example tasks mentioned are shadowing, role-play prompts (e.g., persuasive and interpersonal situation), and hypothetical-situation tasks which are meant to produce extended turns. The technology-mediated support (e.g., shared task sheets and audio recording to self-reflect) was applied to ensure that the rehearsal opportunities and pressure on performance were increased during the first exposure and that performance was not under pressure during initial exposure, which was backed by the research that the appropriate design of the flipped learning could support engagement and performance when in-class time was devoted to the interactive practice instead of lecture (Chang, 2023).The control group adhered to the traditional curriculum/traditional instruction throughout the whole 12-week period and did not get the IFSA It is defined in the , paper as a traditional/conventional teaching (with the references to Grammar-Translation and/or Audio-Lingual orientations) and the control group did not receive any special treatment based on IFSA.

2.3 Instruments

In order to obtain a multidimensional image of the oral proficiency, the research was using three formats of speaking tasks in pretest and posttest: Dialogue Completion Tasks (DCTs), when learners were expecting to talk in a conversational turn with a partner without preparation; Picture-Cued Tasks (PCTs), when the learners were watching an image during two minutes and then were expected to talk about it or to answer questions posed by the interviewer; and Responsive Assessment Tasks (RATs), when the

learners were conversing through a pre-interview The choice of these types of tasks was based on the necessity to sample the different speaking requirements such as prompted turn completion, descriptive talk, and spontaneous responsive talk, as well as to achieve comparability across testing occasions.Fluency/latency and analytic rubric scores.The rating of each activity was based on a 5-point analytic rubric that encompassed five parameters, including time taken to respond (latency/spontaneity), fluency, accuracy, vocabulary and grammar. Latency Response latency was operationalized by explicit timing bands overlaying rubric points: a 5-point score corresponded to an immediate response (less than 2 seconds), 4-point score corresponded to a 3-5-second response, 3-point score corresponded to a 6-8-second response, 2-point score corresponded to a 9-10-second response, and 1-point score corresponded to a response longer than 10 seconds. The rubric therefore considered hesitation/latency as a measurable performance scale, which was used as a behavioral proxy to inhibition in speaking.

The speaking performances were scored by two raters who were trained. The , paper mentions that raters were oriented to the rubric and went through a calibration exercise with anonymized audio samples in the pretest in order to make interpretations of score bands consistent. A subset of the data (20% of assessments) was analyzed on the basis of inter-rater reliability through Cohen kappa, with results of between.78 and.86 (substantial to near-perfect agreement). Any discrepancies greater than one point were addressed by discussion in which consensus scoring was arrived at.

2.3.1 Pre-intervention needs analysis questionnaire.
The researcher created a structured needs-analysis questionnaire that was presented online as a questionnaire on Google Forms before the intervention. The instrument comprised 17 closed and Likert-scale questions that were aimed at collecting data regarding the English speaking patterns of the students, their level of confidence and perceived obstacles, as well as their openness to the instructional process. Questions 9-14 were directly aimed at addressing the obstacles to speaking English (e.g., anxieties and fears and hesitation), whereas questions 15-17 measured the readiness of the learners to the intervention.

2.3.1 Open-ended responses (post-intervention) and feedback questionnaire.
The experimental group was asked to fill out a feedback questionnaire on the basis of a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree) after the intervention. The parts described in the , paper include perceived effectiveness of the flipped component (Items 1-5), immersion/flow experience during the tasks (Items 6-10) as well as perceived emotional/behavioral changes post intervention (including items about reduced fear/anxiety and more direct responding without hesitation). Moreover, qualitative reflections were collected with the help of four open-ended questions

2.3.2 Procedure
The four phases of data collection were carried out. To begin with, it is necessary to have the needs-analysis questionnaire, which would record the baseline learner perceptions about the speaking habits, confidence, barriers, and the willingness to the approach. Second, the analytic rubric was used to score baseline (pretest) speaking performance evoked with the DCT, PCT, and RAT tasks. Third, the experimental group was provided with the 12-week IFSA intervention (36 sessions; three 60-minute sessions per week) and the control group was provided with the standard instruction during the same period. Fourth, the speaking posttest was conducted in the same three task formats and scoring rubric. The experimental group took the feedback questionnaire after posttesting and qualitative responses were obtained through open-ended questions in the context of quantitative results (i.e., changes in anxiety, confidence, and willingness to participate, which were perceived to be) changing).

Data Analysis
The quantitative data were reported to Microsoft Excel where they were entered and cleaned before entering into inferential analysis. The processed dataset was saved in SPSS (Version 28.0 is claimed) where it was processed statistically. Descriptive statistics (means, standard deviations and frequency distributions) of the speaking scores and questionnaire items were used as the starting point of the analysis. Before performing inferential testing, the data was filtered to eliminate

errors in entry, missing data points and outliers; incorrect entries were triangulated against original score sheets and missing data was treated in a conservative fashion to avoid distortion (a single imputation method is not explicitly stated in the , paper). The Shapiro-Wilk test and visual inspection method (Q-Q plots and histograms) were used to determine the test of normality. Paired-samples t tests were conducted to measure change over time in the groups on mean scores at each dimension of the rubric between pretest and posttest, in each of the three speaking task types. Independent-samples t tests were applied in order to compare performance across groups as indicated in the , paper. In the case of the feedback questionnaire on the experimental group, one sample t tests were performed to see whether there were significant mean ratings that were considered significantly above the mean of the neutral point at 3.00. Four Pearson correlation analyses were also done to find out the relationship between pretest and posttest variables. Speaking scores were estimated using Cohen kappa on 20 percent of the assessment sample with all the coefficients of kappa falling within the .78 -.86 range. Descriptive content analysis was used to analyze qualitative data of open-ended feedback questions. In the case of reporting, the , paper reveals that 10 representative responses have been intentionally chosen in each open-ended question to cover thematic coverage without repeating the same. These qualitative results were triangulated and interpreted into the quantitative patterns, especially regarding perceived anxiety reduction, confidence growth, and engagement in speaking tasks, which is of interest in working on the intervention with the focus on affective constraints (Xiong and Zhang, 2024).

The performance on speaking was assessed in the experimental group (Immersive Flow State Approach; IFSA; n = 150) and the control group (n = 50) using three task formats (DCT, PCA, RAT) graded on a 5-point analytic rubric on five dimensions time taken to respond, fluency, relevancy, vocabulary, and grammar. Table 1 will include pretest and posttest descriptive statistics of each group, task, and dimension. The experimental group proved significant improvement in pretest- posttest in all the dimensions in all three types of tasks. As an illustration, the means of the experimental-group have increased in time taken to respond (higher scores indicating faster responding), with DCT increasing in means of 1.40 (SD = 0.50) to 2.88 (SD = 0.53) and in PCA and RAT with 1.28 (SD = 0.46) to 3.16 (SD = 0.55). Similar improvements were also made on fluency (e.g., DCT: 1.32 [SD = 0.48] to 2.96 [SD = 0.51]) and on relevancy (e.g., PCA: 1.20 [SD = 0.41] to 2.96 [SD = 0.54]). There were also significant improvement of vocabulary and grammar scores in tasks of the experimental group. Conversely, control group demonstrated relatively a restricted range of change between pretest and posttest in the majority of different task formats and dimensions. Descriptively, the most significant change that was observed in the control group was in the time taken to respond to the PCA task (1.20 [SD = 0.41] to 1.73 [SD = 0.46]) but most of the other results displayed minuscule shifts (some of them slight decreases, such as fluency on DCT). Table 2 summarizes inferential results of paired samples t tests. All pretest post-test differences of all five rubric dimensions between DCT, PCA, and RAT in the experimental group were found to be statistically significant (reported in , paper as $p = 0.000$ and presented here in APA style as $p < .001$), ranging in t values between 5.527 and 16.830. The PCA of time taken to respond was the only comparison that was statistically significant in the control group, $t = 3.228$, $p = .006$; the rest of the comparisons were not significant ($p > .05$). Further findings on the performance were supported by the results on the learner-perception. In the post-intervention questionnaire of the experimental group, high percentages of respondents agreed and strongly agreed that they have conquered fear and anxiety, they can respond without hesitating, they are more confident to express their thoughts in English, and large proportions of respondents (96, 139, 136) agreed and strongly agreed that they had reduced inhibition and enhanced confidence. Means of the items were substantially greater than the neutral mean (3.00) as were including overcoming fear and anxiety ($M = 4.17$, $SD = 0.83$, one-sample $t = 15.77$, $p < .001$) and responding without hesitation ($M = 4.03$, $SD = 0.89$, one-sample $t = 12.82$, $p < .001$). The qualitative responses were consistent with these tendencies as the students reported having less fear and quicker reaction to supportive activities, interactive practice, role-playing, and being able to learn at their speed with the help of videos.

© FEB 2026 | IRE Journals | Volume 9 Issue 8 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I8-1712260

Table 1

Pretest and Posttest Speaking Performance by Group, Task, and Dimension (Rubric 1–5)

Dimension	Task	Control Pre M (SD)	Control Post M (SD)	Experimental Pre M (SD)	Experimental Post M (SD)
Time taken to respond	DCT	1.73 (0.961)	1.67 (0.488)	1.40 (0.500)	2.88 (0.526)
	PCA	1.20 (0.414)	1.73 (0.458)	1.28 (0.458)	3.16 (0.554)
	RAT	1.27 (0.458)	1.40 (0.507)	1.28 (0.458)	3.16 (0.554)
Fluency	DCT	1.73 (0.594)	1.53 (0.516)	1.32 (0.476)	2.96 (0.611)
	PCA	1.33 (0.488)	1.47 (0.516)	1.40 (0.500)	2.76 (0.663)
	RAT	1.27 (0.458)	1.33 (0.488)	1.44 (0.507)	2.96 (0.539)
Relevancy	DCT	1.53 (0.516)	1.53 (0.516)	1.20 (0.408)	2.80 (0.577)
	PCA	1.40 (0.507)	1.47 (0.516)	1.20 (0.408)	2.96 (0.539)
	RAT	1.33 (0.488)	1.60 (0.507)	1.52 (0.586)	2.92 (0.572)
Vocabulary	DCT	1.47 (0.640)	1.40 (0.507)	1.28 (0.458)	2.32 (0.557)
	PCA	1.27 (0.458)	1.60 (0.507)	1.36 (0.490)	2.24 (0.597)
	RAT	1.20 (0.414)	1.33 (0.488)	1.28 (0.458)	2.60 (0.500)
Grammar	DCT	1.47 (0.516)	1.67 (0.488)	1.29 (0.458)	2.37 (0.557)
	PCA	1.20 (0.414)	1.47 (0.640)	1.36 (0.490)	2.74 (0.597)
	RAT	1.27 (0.458)	1.40 (0.507)	1.28 (0.458)	2.86 (0.500)

Table 2

Paired-Samples t Tests for Pretest–Posttest Change Within Each Group

Dimension	Task	Control *t*	Control *p*	Experimental *t*	Experimental *p*
Time taken to respond	DCT	0.211	.836	12.629	< .001
	PCA	3.228	.006	14.120	< .001
	RAT	0.807	.433	15.667	< .001
Fluency	DCT	0.211	.836	10.830	< .001
	PCA	1.000	.334	9.714	< .001
	RAT	0.323	.751	14.905	< .001
Relevancy	DCT	0.000	1.000	10.474	< .001
	PCA	0.323	.751	16.830	< .001
	RAT	1.740	.104	10.844	< .001
Vocabulary	DCT	0.367	.719	5.527	< .001
	PCA	2.092	.055	6.608	< .001
	RAT	1.000	.334	11.854	< .001
Grammar	DCT	1.000	.334	7.695	< .001
	PCA	1.169	.262	7.695	< .001
	RAT	0.807	.433	10.474	< .001

Note. For results reported as $p = 0.000$ in the , paper, APA reporting conventions are used here ($p < .001$).

Data Analysis

The quantitative data were reported to Microsoft Excel where they were entered and cleaned before entering into inferential analysis. The processed dataset was saved in SPSS (Version 28.0 is claimed) where it was processed statistically. Descriptive statistics (means,

standard deviations and frequency distributions) of the speaking scores and questionnaire items were used as the starting point of the analysis. Before performing inferential testing, the data was filtered to eliminate errors in entry, missing data points and outliers; incorrect entries were triangulated against original score sheets and missing data was treated in a conservative fashion to avoid distortion (a single imputation method is not explicitly stated in the , paper). The Shapiro-Wilk test and visual inspection method (Q-Q plots and histograms) were used to determine the test of normality. Paired-samples t tests were conducted to measure change over time in the groups on mean scores at each dimension of the rubric between pretest and posttest, in each of the three speaking task types. Independent-samples t tests were applied in order to compare performance across groups as indicated in the , paper. In the case of the feedback questionnaire on the experimental group, one sample t tests were performed to see whether there were significant mean ratings that were considered significantly above the mean of the neutral point at 3.00. Four Pearson correlation analyses were also done to find out the relationship between pretest and posttest variables. Speaking scores were estimated using Cohen kappa on 20 percent of the assessment sample with all the coefficients of kappa falling within the .78 -.86 range. Descriptive content analysis was used to analyze qualitative data of open-ended feedback questions. In the case of reporting, the , paper reveals that 10 representative responses have been intentionally chosen in each open-ended question to cover thematic coverage without repeating the same. These qualitative results were triangulated and interpreted into the quantitative patterns, especially regarding perceived anxiety reduction, confidence growth, and engagement in speaking tasks, which is of interest in working on the intervention with the focus on affective constraints (Xiong and Zhang, 2024).

III. RESULTS

The performance on speaking was assessed in the experimental group (Immersive Flow State Approach; IFSA; n = 150) and the control group (n = 50) using three task formats (DCT, PCA, RAT) graded on a 5-point analytic rubric on five dimensions time taken to respond, fluency, relevancy, vocabulary, and grammar. Table 1 will include pretest and posttest descriptive statistics of each group, task, and dimension. The experimental group proved significant improvement in pretest- posttest in all the dimensions in all three types of tasks. As an illustration, the means of the experimental-group have increased in time taken to respond (higher scores indicating faster responding), with DCT increasing in means of 1.40 (SD = 0.50) to 2.88 (SD = 0.53) and in PCA and RAT with 1.28 (SD = 0.46) to 3.16 (SD = 0.55). Similar improvements were also made on fluency (e.g., DCT: 1.32 [SD = 0.48] to 2.96 [SD = 0.51]) and on relevancy (e.g., PCA: 1.20 [SD = 0.41] to 2.96 [SD = 0.54]). There were also significant improvement of vocabulary and grammar scores in tasks of the experimental group. Conversely, control group demonstrated relatively a restricted range of change between pretest and posttest in the majority of different task formats and dimensions. Descriptively, the most significant change that was observed in the control group was in the time taken to respond to the PCA task (1.20 [SD = 0.41] to 1.73 [SD = 0.46]) but most of the other results displayed minuscule shifts (some of them slight decreases, such as fluency on DCT).

Table 2 summarizes inferential results of paired samples t tests. All pretest post-test differences of all five rubric dimensions between DCT, PCA, and RAT in the experimental group were found to be statistically significant (reported in , paper as $p = 0.000$ and presented here in APA style as $p < .001$), ranging in t values between 5.527 and 16.830. The PCA of time taken to respond was the only comparison that was statistically significant in the control group, $t = 3.228$, $p = .006$; the rest of the comparisons were not significant ($p > .05$). Further findings on the performance were supported by the results on the learner-perception. In the post-intervention questionnaire of the experimental group, high percentages of respondents agreed and strongly agreed that they have conquered fear and anxiety, they can respond without hesitating, they are more confident to express their thoughts in English, and large proportions of respondents (96, 139, 136) agreed and strongly agreed that they had reduced inhibition and enhanced confidence. Means of the items were substantially greater than the neutral mean (3.00) as were including overcoming fear and anxiety ($M = 4.17$, $SD = 0.83$, one-sample $t = 15.77$, $p < .001$) and

responding without hesitation (M = 4.03, SD = 0.89, one-sample t = 12.82, p <.001). The qualitative responses were consistent with these tendencies as the students reported having less fear and quicker reaction to supportive activities, interactive practice, role-playing, and being able to learn at their speed with the help of videos.

5.1 Directions for future research

Data Analysis

The quantitative data were reported to Microsoft Excel where they were entered and cleaned before entering into inferential analysis. The processed dataset was saved in SPSS (Version 28.0 is claimed) where it was processed statistically. Descriptive statistics (means, standard deviations and frequency distributions) of the speaking scores and questionnaire items were used as the starting point of the analysis. Before performing inferential testing, the data was filtered to eliminate errors in entry, missing data points and outliers; incorrect entries were triangulated against original score sheets and missing data was treated in a conservative fashion to avoid distortion (a single imputation method is not explicitly stated in the , paper). The Shapiro-Wilk test and visual inspection method (Q-Q plots and histograms) were used to determine the test of normality. Paired-samples t tests were conducted to measure change over time in the groups on mean scores at each dimension of the rubric between pretest and posttest, in each of the three speaking task types. Independent-samples t tests were applied in order to compare performance across groups as indicated in the , paper. In the case of the feedback questionnaire on the experimental group, one sample t tests were performed to see whether there were significant mean ratings that were considered significantly above the mean of the neutral point at 3.00. Four Pearson correlation analyses were also done to find out the relationship between pretest and posttest variables. Speaking scores were estimated using Cohen kappa on 20 percent of the assessment sample with all the coefficients of kappa falling within the .78 -.86 range. Descriptive content analysis was used to analyze qualitative data of open-ended feedback questions. In the case of reporting, the , paper reveals that 10 representative responses have been intentionally chosen in each open-ended question to cover thematic coverage without repeating the same. These qualitative results were triangulated and interpreted into the quantitative patterns, especially regarding perceived anxiety reduction, confidence growth, and engagement in speaking tasks, which is of interest in working on the intervention with the focus on affective constraints (Xiong and Zhang, 2024).

The performance on speaking was assessed in the experimental group (Immersive Flow State Approach; IFSA; n = 150) and the control group (n = 50) using three task formats (DCT, PCA, RAT) graded on a 5-point analytic rubric on five dimensions time taken to respond, fluency, relevancy, vocabulary, and grammar. Table 1 will include pretest and posttest descriptive statistics of each group, task, and dimension. The experimental group proved significant improvement in pretest- posttest in all the dimensions in all three types of tasks. As an illustration, the means of the experimental group have increased in time taken to respond (higher scores indicating faster responding), with DCT increasing in means of 1.40 (SD = 0.50) to 2.88 (SD = 0.53) and in PCA and RAT with 1.28 (SD = 0.46) to 3.16 (SD = 0.55). Similar improvements were also made on fluency (e.g., DCT: 1.32 [SD = 0.48] to 2.96 [SD = 0.51]) and on relevancy (e.g., PCA: 1.20 [SD = 0.41] to 2.96 [SD = 0.54]). There were also significant improvement of vocabulary and grammar scores in tasks of the experimental group. Conversely, control group demonstrated relatively a restricted range of change between pretest and post-test in the majority of different task formats and dimensions. Descriptively, the most significant change that was observed in the control group was in the time taken to respond to the PCA task (1.20 [SD = 0.41] to 1.73 [SD = 0.46]) but most of the other results displayed minuscule shifts (some of them slight decreases, such as fluency on DCT).

Table 2 summarizes inferential results of paired samples t tests. All pretest post-test differences of all five rubric dimensions between DCT, PCA, and RAT in the experimental group were found to be statistically significant (reported in , paper as p = 0.000 and presented here in APA style as p <.001), ranging in t values between 5.527 and 16.830. The PCA of time taken to respond was the only comparison that was statistically significant in the control group, t = 3.228, p =.006; the rest of the comparisons were not significant (p >.05). Further findings on the

performance were supported by the results on the learner-perception. In the post-intervention questionnaire of the experimental group, high percentages of respondents agreed and strongly agreed that they have conquered fear and anxiety, they can respond without hesitating, they are more confident to express their thoughts in English, and large proportions of respondents (96, 139, 136) agreed and strongly agreed that they had reduced inhibition and enhanced confidence. Means of the items were substantially greater than the neutral mean (3.00) as were including overcoming fear and anxiety ($M = 4.17$, $SD = 0.83$, one-sample $t = 15.77$, $p < .001$) and responding without hesitation ($M = 4.03$, $SD = 0.89$, one-sample $t = 12.82$, $p < .001$). The qualitative responses were consistent with these tendencies as the students reported having less fear and quicker reaction to supportive activities, interactive practice, role-playing, and being able to learn at their speed with the help of videos.

REFERENCES

[1] Botes, E., Dewaele, J.-M., & Greiff, S. (2020). The foreign language classroom anxiety scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *Journal of Psycholinguistic Research*.

[2] Chang, Y. (2023). Flipping EFL low-proficiency students' learning: An empirical study. *Language Teaching Research*. https://doi.org/10.1177/13621688231165474

[3] Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

[4] Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132.

[5] Jiang, Y., & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System, 82*, 13–25. https://doi.org/10.1016/j.system.2019.02.017

[6] Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

[7] Payant, C., & Zuniga, M. (2022). Learners' flow experience during peer revision in a virtual writing course during the global pandemic. *System, 105*, 102715. https://doi.org/10.1016/j.system.2021.102654

[8] Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition, 41*(2), 363–387. https://doi.org/10.1017/S027226311800011

[9] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

[10] Xiong, X., & Zhang, L. J. (2024). A meta-analysis and systematic review of foreign language anxiety interventions. *Journal of Language and Social Psychology*. https://doi.org/10.1177/0261927X241291258

[11] Zúñiga, M. (2023). The correlates of flow in the L2 classroom: Linking basic L2 task features to learner flow experiences. *The Modern Language Journal, 107*(3). https://doi.org/10.1111/modl.12865