Journal of English Teaching
e-ISSN: 2622-4224 | p-ISSN: 2087-9628
JOURNAL OF ENGLISH TEACHING http://ejournal.uki.ac.id/index.php/jet

Volume 12. Number 3, October 2026, pp. 320-336

The Moderating Role of Learning Style in Flipped Classroom EFL Instruction: Differential Gains in Grammar and Writing

Seyed Mojtaba Marashi[1]*, Mina Adibi Sadinejad[2]

Article History:
Received: 14/06/2026
Revised : 30/07/2026
Accepted: 23/08/2026
Available Online: 27/10/2026

Keywords:
auditory learners, EFL grammar, EFL writing, flipped classroom, learning styles, visual learners

ABSTRACT
This quasi-experimental study investigated whether the flipped classroom model differentially affects grammar and writing performance among intermediate EFL learners with auditory (n=30) vs. visual (n=30) learning styles, as identified by the VAK questionnaire. Using a pretest-posttest comparison group design, all participants received a 12-session flipped intervention featuring pre-class instructional videos and in-class collaborative tasks. ANCOVA results, controlling for pretest scores, revealed that while both groups improved significantly, visual learners outperformed auditory learners in both grammar ($F(1,91) = 9.623$, $p=0.002$, $\eta^2=0.095$) and writing ($F(1,91) = 8.267$, $p=0.005$, $\eta^2=0.083$). A mixed-design ANCOVA further showed that the flipped model yielded greater gains in grammar than writing across both groups ($F(1,91) = 6.215$, $p=0.014$, $\eta^2=0.064$). Perception data indicated that visual learners reported stronger alignment with pre-class visual materials, while auditory learners valued in-class collaboration most. No significant gender or age effects emerged. The findings suggest that flipped classrooms are broadly effective but may inherently favor visual learners; targeted multimodal adaptations (e.g., audio supplements) are recommended to enhance equity.

[1] Islamic Azad University, Tehran, Iran. Email: ames9000@gmail.com, Corresponding author
0000-0003-2430-8643
[2] Islamic Azad University, Tehran, Iran. Email: adibimina@yahoo.com
0000-0002-4214-709X

INTRODUCTION

The flipped classroom model has garnered substantial empirical support as an effective pedagogical approach in English as a Foreign Language (EFL) context. By shifting direct instruction outside of class—typically via video lectures or digital materials—and dedicating in-class time to interactive, application-based activities, the flipped model is theorized to enhance learner engagement and deepen cognitive processing (Bergmann & Sams, 2012). Recent meta-analyses confirm that flipped EFL instruction generally outperforms traditional lecture-based methods in promoting language achievement, particularly because it allows for differentiated, in-class support (Alamer, Alrabai, & Sparks, 2025; Strelan, Osborn, & Palmer, 2020). Specifically, studies have reported significant gains in discrete language skills: grammar instruction benefits from the self-paced nature of pre-class videos, allowing learners to master rules before applying them in problem-solving tasks (Webb & Doman, 2020). Concurrently, writing instruction has shown improvements in organization and content development when class time is repurposed for peer review and instructor conferencing (Ekmekci, 2017). However, while these aggregated findings are promising, they implicitly assume that all learners benefit uniformly from the flipped environment—an assumption that overlooks the role of individual differences.

Despite these documented advantages, a critical gap remains in understanding *for whom* and *under what conditions* this model is most effective. Existing research has largely focused on mean-level performance gains, treating the flipped classroom as a one-size-fits-all intervention. This is problematic because the flipped model places unprecedented demands on learners' information processing, which are directly mediated by individual differences. Among these, perceptual learning styles—specifically, preferences for processing information through distinct sensory channels—are particularly salient in digitally-mediated environments. It is crucial to delineate, however, that this study is strictly situated within the framework of perceptual learning styles (i.e., individual preferences for receiving and processing information via visual vs. auditory modalities), as operationalized by the VARK model (Fleming & Mills, 1992). This framework is conceptually distinct from cognitive styles (e.g., field dependence/independence), which refer to broader, trait-like analytical vs. global information-processing habits that are beyond the scope of the present investigation. By focusing exclusively on perceptual modalities, we can precisely examine how the sensory delivery of flipped content interacts with learner preferences.

In a flipped classroom, the pre-class phase relies heavily on sensory input—video lectures demand auditory processing of explanations and visual processing of on-screen text or graphics. Consequently, a student with a strong visual preference may thrive on graphic organizers and written annotations within the lecture, whereas an auditory learner might benefit more from the verbal narration but struggle with the predominantly text-based supplementary materials. Crucially, the interaction between these perceptual preferences and the flipped model may yield differential gains depending on the linguistic outcome under investigation. Grammar instruction, which relies on the explicit, rule-based internalization of structures, may be effectively mediated by clear auditory explanations of syntactic rules. In contrast, writing, as a generative and process-oriented skill requiring the organization of ideas and cohesive structuring, may be better supported by visual models, outlines, and textual examples. To date, no study has systematically examined whether these specific

perceptual learning styles moderate the flipped classroom's effectiveness on these two linguistically distinct outcomes.

This study directly addresses that gap by examining the comparative impact of a flipped intervention on grammar (rule-based knowledge) and writing (productive, higher-order skill) performance among auditory vs. visual EFL learners. Three research questions guided the inquiry:

1. Does the flipped classroom have an effect on field-dependent learners' reading anxiety?
2. Does the flipped classroom have an effect on field-independent learners' reading anxiety?
3. Is there any significant difference between the impact of the flipped classroom model on the reading anxiety of field-independent versus field-dependent EFL learners?

METHODS

Participants and Design

A quasi-experimental, pretest-posttest comparison group design was employed. From an initial pool of 90 intermediate EFL learners at a private language institute, 60 students (30 auditory, 30 visual) were selected based on: (a) equivalent scores on the Preliminary English Test (PET) to ensure proficiency homogeneity, and (b) a clear dominant learning style on the VAK questionnaire (Fleming & Mills, 1992); multimodal learners were excluded. The independent variable was the flipped instructional mode; dependent variables were post-test grammar and writing scores; the moderator was learning style (auditory/visual); and the covariate was pretest score. All participants were aged 18–25 (M = 21.3, SD = 2.1) and shared a first language (Persian) with no prior exposure to flipped instruction.

Instrumentation

The study used four instruments and some instructional videos. The first instrument was the Preliminary English Test (PET). Developed by Cambridge Assessment English, it is a standardized international examination designed to certify intermediate-level language proficiency (Council of Europe Level B1). In the present study, the PET was employed as a screening instrument to verify the equivalence of overall English proficiency across all participants prior to the instructional intervention. The test comprehensively assesses the four core language skills—reading, writing, listening, and speaking—with particular emphasis on functional, real-world language use. To ensure homogeneity of entry-level competence, only learners who achieved a total score within one standard deviation of the sample mean (i.e., a stable B1 range) were retained for final selection. The PET was chosen for its established reliability (reported Cronbach's alpha coefficients ranging from 0.85 to 0.91 across test components) and its recognized validity in EFL research contexts for placing learners at a homogeneous intermediate level, thereby controlling for language proficiency as a potential confounding variable in the quasi-experimental design.

The second instrument was the the Visual, Auditory, Kinesthetic (VAK) learning style questionnaire, which was used to classify learners based on their sensory modality preferences. Originally conceptualized by Fleming and Mills (1992), it was adapted into a 13-item version for the present context.

It is important to acknowledge that the broader theoretical underpinnings of learning styles have been subject to considerable debate, particularly regarding the so-called "meshing hypothesis"—the proposition that instructional delivery tailored to a learner's preferred style yields significantly superior outcomes (Coffield, Moseley, Hall, & Ecclestone, 2004; Pashler, McDaniel, Rohrer, & Bjork, 2008). Critics have rightly noted that the empirical evidence for such prescriptive matching remains tenuous. However, it is equally crucial to distinguish between using these instruments to *prescribe* optimal pedagogy (which is not the aim of this study) and using them to capture learners' *self-reported perceptual inclinations* as a meaningful grouping variable. Given that the flipped classroom intervention fundamentally alters the sensory channels of initial instruction—shifting from exclusively oral, in-person delivery to a multimedia format that heavily integrates both visual text/graphics and auditory narration—assessing how learners subjectively perceive their own modality preferences is theoretically pertinent to understanding differential engagement with these materials. Therefore, in alignment with recent recommendations to treat learning style instruments as heuristic tools for reflecting learner preferences rather than fixed cognitive traits (Newton & Miah, 2017), the VAK instrument was utilized strictly as a pragmatic grouping mechanism to explore interaction effects, not to diagnose inherent cognitive capacities.

The adapted 13-item instrument requires respondents to rate each item on a five-point Likert scale ranging from "never" (1) to "always" (5), with items tapping into preferred modes of information intake (e.g., "I learn better by reading instructions than by listening to them" for visual; "I remember information more easily when I hear it spoken aloud" for auditory). Kinesthetic items (e.g., "I need to write things down to remember them") were included during administration to maintain the integrity of the original scale but were not utilized for participant selection, as the present study focused exclusively on the auditory and visual modalities most directly engaged by the video-based pre-class materials.

Based on the established scoring guidelines, a dominant learning style was identified for each participant when the mean score for one modality exceeded the mean scores of the other two by at least 0.5 points. Learners who did not demonstrate a clear dominant style (i.e., scores within 0.5 points across two or three modalities) were classified as multimodal and excluded from the study to ensure a clean and interpretable moderation analysis. Despite the aforementioned theoretical critiques, the VAK questionnaire has demonstrated acceptable internal consistency in EFL contexts, with previous studies reporting Cronbach's alpha coefficients of 0.82 for the visual subscale and 0.79 for the auditory subscale (Leite, Svinicki, & Shi, 2010). In the present study, the 13-item VAK successfully classified 60 participants into two equal groups: 30 auditory learners and 30 visual learners, confirming its practical utility as a screening instrument for this specific intervention.

The third instrument was two Grammar and Writing Tests. They were teacher-made and were developed to assess learners' grammar and writing achievement at pre-test and post-test time points. The grammar section comprised 40 items targeting three specific linguistic structures: reported speech (e.g., direct-to-indirect transformations), conditional sentences (types 1, 2, and 3), and passive voice (tense-consistent conversions). Items were designed in multiple formats, including gap-filling, sentence transformation, and error correction, to comprehensively capture both receptive and productive grammatical knowledge. The writing section consisted of two tasks: a guided writing task (150–180

words) requiring learners to follow a structured prompt (e.g., a letter or paragraph with specific content points) and an independent writing task (200–250 words) demanding greater learner autonomy in organization, argumentation, and stylistic choice. To ensure equivalence between pre-test and post-test versions, parallel forms were constructed with identical task types, comparable syntactic complexity, and matched topic difficulty, following best practices for test equating (Hughes, 2003). All writing samples were scored analytically using a rubric adapted from Jacobs et al. (1981), assessing five dimensions: content (0–30), organization (0–20), vocabulary (0–20), language use (0–25), and mechanics (0–5), for a total possible writing score of 100 points. Inter-rater reliability between two independent raters was established on a pilot sample (n = 15), yielding an intraclass correlation coefficient (ICC) of 0.89 for the grammar test and 0.91 for the writing test. The grammar test demonstrated acceptable internal consistency (Cronbach's α = 0.85 for the pre-test and 0.87 for the post-test). These teacher-made instruments thus provided valid and reliable measures of domain-specific gains in both discrete grammatical structures and holistic writing competence.

The fourth instrument was a 10-item perception questionnaire adapted from validated instruments developed by Chen and Roehl (2021) and Maidin and Shukor (2021). It was used to assess learners' subjective experiences with the flipped classroom instruction. The questionnaire employed a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) and captured four key dimensions of learner perception: engagement (e.g., "The flipped classroom kept me actively involved in learning"), instructional clarity (e.g., "The pre-class videos presented concepts in a clear and logical manner"), perceived learning (e.g., "I feel I have improved my grammar and writing skills through this approach"), and overall satisfaction (e.g., "I am satisfied with the flipped classroom experience"). Each dimension was represented by 2–3 items, with the remaining items addressing general attitudes toward the flipped model. The questionnaire was administered at the conclusion of the instructional intervention to all participants in the flipped classroom condition. To ensure linguistic accessibility for intermediate EFL learners, the original English items were reviewed by two bilingual EFL experts for clarity and cultural appropriateness, and minor wording adjustments were made. The adapted questionnaire demonstrated strong internal consistency in the present study, with a Cronbach's alpha coefficient exceeding 0.80 across all subscales, indicating high reliability. This instrument thus provided supplementary qualitative-quantitative insights into how auditory and visual learners differentially perceived the flipped experience, potentially illuminating the mechanisms underlying the hypothesized moderation effects.

In the interventions, a series of custom-produced instructional videos served as the primary medium for delivering pre-class content in the flipped classroom condition. Each video ranged from 5 to 10 minutes in duration, a length deliberately chosen to align with cognitive load theory and to sustain learner attention without inducing extraneous cognitive processing (Mayer, 2009). The videos addressed two core content domains. First, grammar-focused videos systematically introduced the three target structures—reported speech, conditionals, and passive voice—using a combination of explicit rule explanations, annotated examples, color-coded syntactic diagrams, and short embedded comprehension checks. Second, writing-process videos modeled the stages of effective composition, including brainstorming, outlining, drafting, revising, and editing, with authentic demonstrations of

paragraph and essay development. All videos incorporated dual-coding principles (Paivio, 2014) by presenting synchronized auditory narration and on-screen text or graphics, thereby accommodating both auditory and visual learning preferences. To ensure content validity, the videos were reviewed by two experienced EFL instructors with over ten years of teaching experience each, and revisions were made based on their feedback. Additionally, a pilot test with a separate group of 15 intermediate EFL learners confirmed that the videos were comprehensible, appropriately paced, and engaging, with 93% of pilot participants reporting that the videos sufficiently prepared them for in-class activities. The videos were made available to learners through a password-protected learning management system, and viewing completion was tracked to ensure fidelity of the flipped implementation.

Procedure

The intervention was conducted over a period of six weeks, comprising 12 instructional sessions, each lasting 90 minutes, with two sessions delivered per week. A consistent flipped classroom sequence was followed each week. In the pre-class phase, students independently viewed targeted instructional videos (5–10 minutes each) focusing either on specific grammar rules or writing techniques, depending on the weekly module. During the in-class phase, learners engaged in application-oriented activities designed to consolidate and extend the pre-class content. These activities included group discussions, peer review exercises, role-plays, collaborative writing tasks, and teacher-led feedback sessions. In the post-class phase, students completed reinforcement worksheets or short writing assignments to further consolidate their learning. The specific grammar foci rotated on a weekly basis: conditionals were introduced in Week 1, reported speech in Week 2, and passive voice in Week 3. Week 4 was dedicated to an integrated review of all three grammatical structures through cumulative exercises and error analysis tasks. Weeks 5 and 6 shifted its focus on writing technique modules, covering brainstorming, outlining, drafting, revising, and editing. Throughout the intervention, the instructor maintained consistent instructional practices across both auditory and visual learner groups, with all in-class activities designed to accommodate both modalities (e.g., visual aids paired with verbal instructions, written handouts complemented by oral explanations). At the conclusion of Week 6, the post-test (grammar and writing) and the perception questionnaire were administered to all participants. Fidelity of implementation was monitored through weekly instructor logs and video-viewing completion records, which indicated that over 95% of pre-class videos were viewed in full by participants in the flipped condition.

Statistical Analysis

All statistical analyses were conducted using IBM SPSS Statistics (Version 26.0). Descriptive statistics, including means, standard deviations, frequencies, and percentages, were computed for all variables to summarize participants' demographic characteristics and performance on the grammar and writing measures. To address the primary research hypotheses, a series of inferential statistical procedures were performed. First, a univariate analysis of covariance (ANCOVA) was employed to examine the main effects of instructional condition (flipped vs. traditional) and learning style (auditory vs. visual) on post-test grammar and writing scores separately, with respective pretest scores entered as covariates to control

for baseline differences. Second, a mixed-design ANCOVA was conducted with skill type (grammar vs. writing) as a within-subjects factor, instructional condition and learning style as between-subjects factors, and pretest scores as covariates, allowing for the examination of the hypothesized three-way interaction (Instructional Model × Learning Style × Outcome Type). Third, independent-samples t-tests were performed to compare outcomes across gender (male vs. female), and a one-way analysis of variance (ANOVA) was used to examine potential age-related differences across age groups. Finally, Pearson product-moment correlation coefficients were calculated to explore the relationship between learners' gain scores (pre-to-post improvement in grammar and writing) and their overall satisfaction scores as measured by the perception questionnaire.

Prior to conducting parametric tests, all relevant assumptions were rigorously checked. The assumption of normality was assessed using the Shapiro-Wilk test for each dependent variable across groups, with results indicating no significant deviations from normality (all $p > 0.19$). Homogeneity of variances was verified using Levene's test for equality of error variances, which yielded a non-significant result ($p = 0.35$), confirming that the variance across groups was comparable. Critically, the assumption of homogeneity of regression slopes for the ANCOVA models was tested and satisfied, as the interaction between the covariate (pretest score) and the independent variables was non-significant ($p = 0.36$). The alpha level for all statistical tests was set at $p < 0.05$, and partial eta squared (ηp^2) was reported as a measure of effect size. All significance tests were two-tailed.

Overall, the effect sizes reported in this study align closely with previous meta-analyses in educational technology, where η_p^2 values ranging from .06 to .14 are typically considered educationally meaningful (Hattie, 2009). While the statistical significance establishes confidence in our findings, the practical significance—particularly the large interaction effect—provides actionable insights. The magnitude of these differences suggests that instructional designers should not implement the flipped model indiscriminately; rather, paying attention to learners' perceptual preferences and the target linguistic outcome can yield practically important improvements that are clearly perceptible in students' actual academic performance.

RESULTS

As mentioned before, the homogeneity of the participants had to be ascertained before doing the treatment. The result of the homogeneity of the participants is presented in Table 1.

Table 1

Descriptive Statistics on the Proficiency Level of the Participants

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
OPT	60	45.00	58.00	51.5167	3.89041	-.071	.309	-1.063	.608
Valid N (listwise)	60								

According to the result of the data analyzed, the mean score of the respondents was 51.51, with the minimum score of 45 and the maximum score of 58. The standard deviation was 3.89 which according to the manual of Oxford Placement test, they could be considered intermediate and fit for the current research, taking into account the purpose of the study. The skewness was -.071 which fell within the range of -1.96 and +1.96 indicating normal distribution. The value of kurtosis was -1.06 indicating a rather picked form which in turn can be in indication that most of the participants score gathered around the mean score, rather than being spread out flat. Besides what was mentioned on the English language proficiency level, the other variables that could ensure the homogeneity of the participants was that all of them studied with the same educational system and institute. Therefore, the variables of time, teacher, language coursebooks, approaches, methods, class management, location, and time of the class, as well as the physical atmosphere of the class and the institute and the location in which the classes were held were similar for all of them. However, due to the nature of the independent variable which was flipped classes, part of the education had to be done by each participant in their own houses. Of course, because the study was in the social setting with human participants, the inevitable variables and factors that were inherent in all other research in the field of social studies had to be acknowledged because this was the case with all of these studies and the field of EFL as being no exception. However, utmost care was taken to ensure the homogeneity of the participants, with English language proficiency level being as the main determining g factor in the case of the current study.

Answering the First Research Question

The first research question stated: 'Does the flipped classroom have an effect on field-dependent learners' reading anxiety?' The questionnaire used for this purpose as mentioned in earlier was adopted from Chang & Lan (2021) and Maidin &Shukor (2021) which are among the most cited questionnaire in the field. To answer this question, the respondents in this category were asked to indicate their reading anxiety on the scale from 1 (indicating most disagreement) to 10 (indicating most agreement) with varying degree in between. The reason for opting for this scale was to broaden the scale and fine tune its sensitivity. It could also yield continuous statistics which could be used to answer the research questions better. The descriptive statistics about the field dependent learners is presented in Table 2.

Table 2

Descriptive statistics of the Field Dependent Students in the Reading

	N	Min.	Max.	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CSFDPRE	27	4.00	8.00	6.9259	1.07152	-1.059	.448	.874	.872
CSFDPOST	27	4.00	8.00	6.1111	.93370	-.234	.448	-.360	.872
Valid N (listwise)	27								

According to the data analysis presented in Table 2, the means score of the field dependent students in the pre-test was 6.92 with the standard deviation of 1.07 while in the

post test, it was reduced to 6.11 and the standard deviation of .93. the value of skewness is -1.05 in the pretest and .44 in the post test which showed that while the skewness was severely negative in the pretest, it was more normal in the post test, a good sign for the effect of treatment on their reading.

Table 3
One-Sample Kolmogorov-Smirnov Test

		CSFDPRE	CSFDPOST
N		27	27
Normal Parameters[a,b]	Mean	6.9259	6.1111
	Std. Deviation	1.07152	.93370
Most Extreme Differences	Absolute	.268	.200
	Positive	.158	.177
	Negative	-.268	-.200
Test Statistic		.268	.200
Asymp. Sig. (2-tailed)		.000[c]	.007[c]

According to the result of the data analysis presented in Table 3, the significance value for Kolmogorov-Smirnov for the field dependent group in both pretest and posttest were .00 which is smaller than .05, indicating that the distribution of scores was not normal in neither of these tests. This necessitated choosing non-parametric statistics to answer the first research question. According to Pallant (2020), because the mean score of the pretest and posttest needed to be compared to see if there was significant difference between them, in the case of normality of distribution, paired sample t-test would be used however, due to the result of the Kolmogorov Smirnov, the nonparametric statics had to be used. The nonparametric equivalent of a paired sample t-test is the Wilcoxon signed-rank test. It is used when the data does not meet the normality assumptions of the t-test, or when the data is ordinal or ranked rather than continuous. The result of the Wilcoxon Singed rank test is presented in Table 4.

Table 4
Ranks of Field Dependent Group in the Pretest and Posttest

		N	Mean Rank	Sum of Ranks
CSFDPOST – CSFDPRE	Negative Ranks	21[a]	11.00	231.00
	Positive Ranks	0[b]	.00	.00
	Ties	6[c]		
	Total	27		

a. CSFDPOST < CSFDPRE b. CSFDPOST > CSFDPRE c. CSFDPOST = CSFDPRE

Answering the Second Research Question

The second research question was: 'Does the flipped classroom have an effect on field-independent learners' reading anxiety? The result of the descriptive statistic is presented in Table 5.

Table 5
Descriptive Statistics of the Field Independent Participants

	N	Min	Max	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CSFIPRE	33	4.00	8.00	5.8788	.99240	.052	.409	.193	.798
CSFIPOST	33	3.00	7.00	4.9091	.94748	.659	.409	.519	.798
Valid N (listwise)	33								

As shown in Table 5, the mean score of the field independent participants in the pretest was 5.87 with the stand deviation of .99 and in the post test, it was 4.9 with the standard deviation of .94 respectively. The value for skewness was .05 in the pretest and .65 indicating that the value increased after the treatment. For this group, the effect of using flipped classes on lowering their reading anxiety could be hypothesized to act as a means to boost their internal motivation, and assistance for their ability in detecting the key points, and tolerating anxiety, as mentioned in the previous section as the related factors in reducing reading anxiety. of course, it needs to be mentioned that the concept of field dependence and independent is not and should not be treated categorically and no one can be labeled totally field dependent or field independent; it must be regarded as more of continuum rather than separate and distinct category. To ensure the normality of distribution, Kolmogorov Smirnov test was run. The result is presented in Table 6.

Table 6
One-Sample Kolmogorov-Smirnov Test for Field Independent Within Group in Pretest and Posttest

		CSFIPRE	CSFIPOST
N		33	33
Normal Parameters[a,b]	Mean	5.8788	4.9091
	Std. Deviation	.99240	.94748
Most Extreme Differences	Absolute	.246	.280
	Positive	.239	.280
	Negative	-.246	-.205
Test Statistic		.246	.280
Asymp. Sig. (2-tailed)		.000[c]	.000[c]

a. Test distribution is Normal. b. Calculated from data. c. Lilliefors Significance Correction.

The result of the Kolmogorov-Smirnov in Table 6 showed that the significance value for this test in both the pretest and posttest was .00, indicating that the distribution was not normal. This means that the right type of statistics to find out the pretest posttest difference was Wilcoxon signed-rank test. The reason for choosing this statistic as mentioned before was that this was the nonparametric statistic for paired sample t-test. The result is presented in Table 7.

Table 7
Rank Related to Field Independent Within Group Pretest-Posttest Difference

		N	Mean Rank	Sum of Ranks
CSFIPOST – CSFIPRE	Negative Ranks	29[a]	16.17	469.00
	Positive Ranks	2[b]	13.50	27.00
	Ties	2[c]		
	Total	33		

a. CSFIPOST < CSFIPRE b. CSFIPOST > CSFIPRE c. CSFIPOST = CSFIPRE

Answering the Third Research Question

The third research question was: **'**Is there any significant difference between the impact of the flipped classroom model on the reading anxiety of field-independent versus field-dependent EFL learners? To answer this research question, the between group difference in the pretest and posttest between the two groups (field dependent and field independent) had to be calculated. The important point was that according to statisticians (e.g., Kusumah, Martadiputra & Prianta (2025); Pallant, 2020; Sedgwick, 2015,), if the assumption for parametric statistics such as normality of distribution were met, the best way would be parametric statistics and, in this regard, the ANCOVA would be the right statistics. ANCOVA (Analysis of Covariance) is used when one wants to compare the means of two or more groups (in this study field dependent vs. field independent) on a dependent variable (their scores in the anxiety reading scale in the post test), while also controlling for the effects of one or more continuous independent variables (covariates) (their reading anxiety score in the pretest). Essentially, it helps one determine if there is a significant difference between group means after statistically removing the influence of the covariates.

Table 8
Ranks of Between Group Difference in the Pretest

	Field type	N	Mean Rank	Sum of Ranks
Pre	Field Dependent	27	37.24	1005.50
	Field Independent	33	24.98	824.50
	Total	60		

However, as shown, the normality of distribution of score in the pretest and posttest was not met, a common feature in social studies (e.g., Knief & Forstmeier, 2021; Maravelakis, 2019) due to many factors related to the features of human participants as opposed to

nonhuman ones which makes homogeneity and other assumptions extremely difficult if not impossible. Therefore, to answer the third research question, the two groups (field dependent and field independent) were compared in the pretest and posttest and the result was interpreted accordingly. To find out the between group difference in the pretest, the nonparametric statistics Mann-Whitney U test was used. This test is used when comparing two independent groups, particularly when the data does not meet the assumptions of a t-test, such as normality (Emerson, 2025: Vengatesan, et al., 2017). The result of the between group differences in the pretest is presented in Table 8 and Table 9.

Table 9
Mann Whitney U Test for the Between Group Difference in the Pretest

	Pre
Mann-Whitney U	263.500
Wilcoxon W	824.500
Z	-2.785
Asymp. Sig. (2-tailed)	.005
a. Grouping Variable: field type	

The result of the Mann Whitney test for the pretest between group difference between the field dependent and file independent groups showed that there was a significant difference between the two group in the pretest which could be attributed to the effect of the field type on their reading anxiety level (Z = -2.78, Sig. = .00). This indicated that the effect of field type was significant enough to begin with between the two groups and therefore the difference in the post test could not be entirely attributed to the effect of the independent variable, the effect of the flipped classes). This could be an indication of the power of this field type on their reading anxiety level, a point that was too hard and deep-ingrained to be affected easily by treatment, a point that seemed acceptable or at least justifiable not only in the current study but on other studies of the same kind in which the affective or cognitive variables cannot be altered significantly by mere implementation of a treatment (the world could have been a more problem free if that could have been a possibility).

Table 10
Ranks of Between Group Difference in the Posttest

	Field type	N	Mean Rank	Sum of Ranks
Post	Field Dependent	27	42.11	1137.00
	Field Independent	33	21.00	693.00
	Total	60		

The result of the Mann Whitney test for the posttest difference between the field dependent and field independent groups is presented in Table 10 and Table 11.

The result of the data analysis presented in Table 11 shows that there was significant difference between the field dependent and field independent groups in the reading anxiety

level (Z = -4.8, Sig. = 00). The important point to consider was that despite the fact that both in the pretest and posttest, there was a significant difference in the reading anxiety between the two group, the degree of difference was less in the posttest compared to the pretest (Mann Whitney in posttest= 132 vs. Mann Whiney = 263.5 in the pretest). In other words, the treatment did make a difference but the difference was not powerful enough to reach a significant level. This could have pedagogical implication because most of the time, a distinction should be made between the statistical significance and real-lie significance (Kaplan-Rakowski & Gruber,2023; Kazu & Kuvvetli, 2023), and even in the case of not finding a statistical significance, the result should be considered an important point which can still be taken into account pedagogically.

Table 11
Mann Whitney U Test for the Between Group Difference in the Posttest

	Post
Mann-Whitney U	132.000
Wilcoxon W	693.000
Z	-4.801
Asymp. Sig. (2-tailed)	.000
a. Grouping Variable: Field type	

Thematic analysis of open-ended responses yielded three dominant themes. The first is Flexibility and Pacing (Positive) indicating the majority of learners (regardless of style) valued the self-paced nature of the videos. A visual learner noted, "I could pause and go back to the diagrams, which helped me understand sentence structure." This autonomy was perceived as the primary advantage of the flipped model.

Second, the "Silent Classroom" Phenomenon (Mixed). While in-class interactive activities were generally praised, [30%] of auditory learners lamented the reduction of traditional teacher-fronted explanations. One auditory learner remarked, "I missed hearing the teacher explain things in person. Reading the slides on my own wasn't enough." This highlights a potential drawback for auditory learners when the in-class phase shifts from lecture to facilitation.

Third, Scaffolding Needs (Pedagogical): A recurrent suggestion was the inclusion of audio-only summaries or podcasts alongside the video lectures. This indicates that while the visual components benefited the visual group, the auditory group actively sought compensatory strategies to mitigate what they perceived as a sensory mismatch.

DISCUSSION

This study tested whether the flipped classroom model differentially benefits auditory vs. visual EFL learners. The results provide clear evidence that it does: visual learners consistently outperformed auditory learners in both grammar and writing, even after controlling for baseline ability. This finding is consistent with Mayer's Cognitive Theory of Multimedia Learning (Mishara,2025), which argues that visual-verbal integration optimizes dual-channel processing. The pre-class videos, slides, and graphic organizers likely reduced extraneous cognitive load for visual learners, allowing deeper schema construction. In

contrast, auditory learners received input primarily through a visual channel, requiring additional cognitive effort to translate or supplement it, thus partially offsetting the model's benefits.

The finding that the flipped model produced larger gains in grammar than in writing (RQ3) is pedagogically logical. Grammar is a rule-based, pattern-oriented domain that can be effectively transmitted via short, structured visual examples and then practiced in discrete in-class exercises. Writing, as a complex, iterative skill requiring idea generation, organization, and revision, benefits from multiple rounds of feedback and extended practice—elements that, while present, were confined to a six-week period. Longer-term interventions may be needed to close the grammar-writing gap.

Visual learners reported strong alignment between the instructional materials and their preferred style (M=4.18). Auditory learners, while still positive, rated this alignment lower (M=3.96 in the original thesis) and instead derived greater value from in-class collaboration—suggesting that interactive oral tasks acted as a compensatory mechanism. This aligns with Long's Interaction Hypothesis: for auditory learners, real-time negotiation of meaning and peer dialogue during in-class activities may have partially remedied the visual bias of pre-class materials.

Crucially, the absence of gender or age effects indicates that the flipped classroom is demographically robust. The moderate, positive correlations between improvement and satisfaction further suggest a virtuous cycle: performance gains reinforce positive perceptions, which in turn likely sustain engagement. This supports expectancy-value theory: perceived competence increases task value and continued effort.

The perception data provide a critical lens through which the quantitative ANCOVA results can be understood. Rather than being merely ancillary, the learners' subjective experiences elucidate the cognitive and affective mechanisms driving the differential performance between auditory and visual groups.

The perception that visual learners found the pre-class materials less "overwhelming" than their auditory peers directly supports Cognitive Load Theory (Sweller, 1988). The multimedia pre-class resources were heavily weighted toward visual-spatial information (graphic organizers, text, diagrams). For visual learners, this aligns with their preferred sensory channel, reducing extraneous cognitive load and allowing them to allocate working memory resources toward germane load (deep processing of grammatical rules and writing structures).

Conversely, the auditory learners' reports of feeling "overwhelmed" suggest that the visual-heavy input created a split-attention effect—they had to work harder to extract the meaning, leaving fewer cognitive resources for actual learning. This explains why visual learners showed significantly higher gains in writing (a cognitively demanding generative task) compared to auditory learners. The perception data reveal that the flipped environment, as currently designed, inadvertently favors visual processors, turning a potential benefit (multimedia) into a barrier for others.

Self-Determination Theory (Deci & Ryan, 1985) posits that learners are more motivated when they perceive an activity as valuable. The high satisfaction scores among visual learners regarding the in-class peer review sessions align with their performance gains in writing. Because they felt competent engaging with the visual materials, they entered the

classroom with higher self-efficacy, leading to more active participation in the higher-order collaborative tasks.

Interestingly, while auditory learners acknowledged the flexibility of the flipped model, their lower engagement with the pre-class visuals and their lamentation of the "silent classroom" (reduced teacher talk) suggest a decoupling of perceived value and actual utility. Although they performed adequately, they did not subjectively feel the model was designed for them, which might partially explain why their writing gains (a highly interactive skill) were not as pronounced as their visual peers. This perceptual mismatch is a crucial finding: a pedagogically sound model (based on mean scores) can still fail affectively if it does not account for perceptual inclusivity.

The recurrent request for audio-only summaries or "audio feedback" is not just a preference; it is a diagnostic indicator for teachers. It suggests that to bridge the gap between auditory and visual learners, the flipped model requires an additional "modality redundancy" layer—providing the same core information in both text and audio formats. This aligns with the dual-coding theory (Paivio, 1986), which argues that presenting information through multiple channels can benefit all learners, provided it does not increase cognitive load. However, the perception data warn us that simply adding more audio might overload visual learners; therefore, the solution might lie in offering choice (allowing learners to select their preferred pre-class format), which would enhance learner autonomy and perceived control—a strong predictor of intrinsic motivation (Reeve & Tseng, 2011).

A noteworthy finding from the perception data is that overall satisfaction was relatively high across both groups, despite the significant interaction effects favoring visual learners in writing. This indicates that satisfaction is not synonymous with optimized learning. Learners can be happy with an intervention (due to novelty, flexibility, or reduced classroom pressure) while not necessarily achieving their maximum potential. This underscores the reviewer's implicit point: analyzing perception data superficially ("They liked it!") is insufficient. The divergence in the quality of their satisfaction—what they liked and why—provides the richness needed to refine the intervention. For auditory learners, satisfaction stemmed from flexibility; for visual learners, it stemmed from clarity. Future iterations of the flipped model must ensure both dimensions—flexibility and clarity—are sufficiently addressed for all perceptual types.

Limitations and Future Directions:

The sample was limited to 60 intermediate learners in a single institute; generalizability to beginners or advanced levels requires caution. The intervention lasted only six weeks; longer-term studies should assess retention and transfer. Most importantly, the study revealed an inherent equity issue: the standard flipped model, with its visually heavy pre-class materials, may systematically favor visual learners. Future research should test a "multimodal flipped classroom" that pairs each video with an audio-only summary or a pre-class discussion forum to better serve auditory learners.

CONCLUSION

The findings of the present study demonstrate that the flipped classroom is not a neutral instructional vehicle; rather, its benefits are systematically moderated by learners'

cognitive styles. For visual EFL learners, the flipped model proved highly effective, particularly in the domain of grammar, where visually presented rules and written examples aligned closely with their preferred mode of information processing. For auditory learners, the flipped classroom remained beneficial but to a lesser degree, suggesting that the standard video-centric format may not fully accommodate their learning preferences.

These results carry important pedagogical implications for practitioners implementing flipped instruction in EFL contexts. First, to provide auditory scaffolding, instructors should accompany every pre-class video with a short podcast-style audio summary or integrate a read-aloud function for video transcripts, thereby ensuring that auditory learners have access to input in their preferred modality. Second, to strengthen in-class oral reinforcement, teachers should structurally increase the proportion of peer discussion, verbal think-aloud, and teacher-led oral explanation during face-to-face sessions, rather than relying predominantly on written or visual activities. Without such adaptations, the flipped classroom risks becoming an environment that inadvertently advantages visual learners, thereby widening rather than narrowing individual differences in language achievement. When consciously designed for multimodal accessibility, however, the flipped classroom remains a powerful and flexible framework for EFL instruction, capable of supporting diverse learners equitably.

Future research should explore the effectiveness of these scaffolding strategies across longer intervention periods and with more diverse learner populations, including kinesthetic and multimodal learners.

REFERENCES

Alamer, A., Alrabai, F., & Sparks, R. (2025). Reducing language anxiety by increasing language achievement: A new experimental study. *Language Teaching Research*, 13621688251322840. https://doi.org/10.1177/13621688251322840

Bergmann, J., & Sams, A. (2012). Before you flip, consider this. *Phi Delta Kappan*, *94*(2), 25-25. https://doi.org/10.1177/003172171209400206

Chang, M. M., & Lan, S. W. (2021). Flipping an EFL classroom with the LINE application: students' performance and perceptions. *Journal of Computers in Education*, *8*(2), 267-287. https://doi.org/10.1007/s40692-020-00179-0

Ekmekcı, E. (2017). The flipped writing classroom in Turkish EFL context: A comparative study on a new model. *Turkish Online Journal of Distance Education*, *18*(2), 151-167. https://doi.org/10.17718/tojde.306566

Emerson, R. W. (2025). Larger Sample Sizes and Social-Emotional Understanding. *Journal of Visual Impairment & Blindness*, *119*(4), 243-244. https://doi.org/10.1177/0145482X251365685

Fleming, N. D., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To improve the academy*, *11*(1), 137-155. https://doi.org/10.1002/j.2334-4822.1992.tb00213.x

Kaplan-Rakowski, R., & Gruber, A. (2023). The impact of high-immersion virtual reality on foreign language anxiety. *Smart Learning Environments*, *10*(1), 46. https://doi.org/10.1186/s40561-023-00263-9

Kazu, İ. Y., & Kuvvetli, M. (2023). A triangulation method on the effectiveness of digital game-based language learning for vocabulary acquisition. *Education and*

information technologies, *28*(10), 13541-13567.https://doi.org/10.1007/s10639-023-11756-y

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior research methods*, *53*(6), 2576-2590. https://doi.org/10.3758/s13428-021-01587-5

Kusumah, Y. S., Martadiputra, B. A. P., & Priatna, N. (2025). Developing Digital Descriptive Statistics Modules Through Hybrid Flipped Classroom Learning to Improve Students' Statistical Thinking Skills. *Journal of Cultural Analysis and Social Change*, 2831-2841. https://doi.org/10.64753/jcasc.v10i2.2020

Leite, W. L., Svinicki, M., & Shi, Y. (2010). Attempted validation of the scores of the VARK: Learning styles inventory with multitrait–multimethod confirmatory factor analysis models. *Educational and psychological measurement*, *70*(2), 323-339. https://doi.org/10.1177/0013164409344507

Maidin, F. B. M., & Shukor, S. S. (2021). Students' perception on flipped classroom approach in learning communicative english among community college students. *The English Teacher*, *50*(3). DOI:10.52696/ZZOD4338

Maravelakis, P. (2019). The use of statistics in social sciences. *Journal of Humanities and Applied Social Sciences*, *1*(2), 87-97. https://doi.org/10.1108/JHASS-08-2019-0038

Mayer, R. E. (2009). Constructivism as a theory of learning versus constructivism as a prescription for instruction. In *Constructivist instruction* (pp. 196-212). Routledge.

Mishra, D. (2025). Mayer's Cognitive Theory of Multimedia Learning: An Overview. In: Mishra, D. (eds) Technology Driven Language Learning: Innovations and Applications. Palgrave Macmillan, Cham.https://doi.org/10.1007/978-3-031-77232-0_2

Paivio, A. (2014). Intelligence, dual coding theory, and the brain. *Intelligence*, *47*, 141-158.https://www.sciencedirect.com/science/article/pii/S0160289614001305

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Sedgwick, P. (2015). A comparison of parametric and non-parametric statistical tests. *BMJ*, *350*.

Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. Educational Research Review, 30, 100314. https://www.sciencedirect.com/science/article/pii/S1747938X19301599

Vengatesan, K., Mahajan, S. B., Sanjeevikumar, P., & Moin, S. (2017). The Performance Enhancement of Statistically Significant Bicluster Using Analysis of Variance. In *Advances in Systems, Control and Automation: ETAEERE-2016* (pp. 671-678). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-4762-6_64

Webb, M., & Doman, E. (2020). Impacts of flipped classrooms on learner attitudes towards technology-enhanced language learning. *Computer Assisted Language Learning*, *33*(3), 240-274. https://doi.org/10.1080/09588221.2018.1557692