

Understanding Classroom Silence: A Cognitive-Decision Model of Reference-Point Dependence and Task Framing



Baoxia Zuo^{1,*}

¹*School of Humanities, Tianjin University of Finance and Economics Pearl River College, P.R. China*

Abstract: Classroom silence persistently constrains college oral English instruction, yet its cognitive mechanisms remain underexplored. Drawing on Prospect Theory, this study constructs and tests an integrated model linking reference-point dependence, task framing, and language proficiency to explain oral risk-aversion. Survey data from 541 non-English-major undergraduates were analyzed via structural equation modeling. Results show that (1) four reference-point types-peer comparison, self-expectation, cultural norms, and teacher specification-significantly predict risk-aversion, with teacher specification exerting the strongest effect ($\beta = 0.876$); (2) perceived teacher task framing partially mediates this relationship (54.7% of total effect); (3) task framing and reference-point dependence interact to form a "cognitive-situational reinforcement loop"; (4) multi-group SEM based on CET-4 passing status reveals that non-passers show stronger direct reference-point→risk-aversion paths, while passers show a stronger task-framing→risk-aversion path; notably, the teacher-specification→risk-aversion path remains stable across proficiency levels. These findings position reference-point dependence as a key cognitive driver of classroom silence and task framing as an actionable intervention channel, while proficiency-stratified sensitivities call for differentiated pedagogical strategies. This study offers a novel cognitive-decision perspective on L2 classroom interaction and provides empirical and theoretical grounds for task and assessment design.

Keywords: foreign language learners, classroom oral performance, risk aversion, reference-point dependence, prospect theory, language proficiency

1. Introduction

Classroom silence-manifested as speech avoidance, simplified expression, or total reticence-has become a persistent bottleneck in college oral English instruction. Even learners with adequate linguistic knowledge often refrain from speaking or resort to low-risk expressions in impromptu oral tasks, seeking to avoid the psychological costs of negative evaluation, public error, or peer pressure. This avoidance not only reduces output frequency and quality but also disrupts the input-interaction-production cycle, trapping oral proficiency in a prolonged "passive cognitive" stage. Breaking this impasse thus requires not simply encouraging students to speak, but understanding the decision-making logic underlying silence.

Existing accounts attribute oral risk-aversion primarily to affective traits (e.g., foreign language anxiety) or social contexts (e.g., face culture, teacher evaluation). Both approaches, however, treat behavior as a direct response to external stimuli, overlooking the active cognitive appraisal that

precedes the decision to speak or remain silent. Prospect Theory offers a complementary lens: it posits that decision-making under uncertainty is governed not by objective probabilities but by subjective reference points, and features systematic loss aversion-the avoidance of loss outweighing the pursuit of equivalent gains. In classroom terms, when learners anticipate that speaking may result in errors, their decision frame shifts to the "loss domain" (face, evaluation, opportunity), triggering risk-avoidant responses. While this perspective opens a promising cognitive-decision pathway to understanding classroom silence, systematic empirical testing within Chinese college foreign language contexts remains scarce.

Moreover, the reference points learners draw upon in their decisions are multi-sourced-peer performance, self-expectations, cultural norms, and teacher-set criteria and task requirements. Among these, the teacher's discursive framing of tasks (gain frame vs. loss frame) is particularly consequential, as it may fundamentally alter learners' risk perception of the same task. Yet extant research has not systematically addressed how multiple reference points jointly predict oral risk-aversion, whether task framing amplifies or buffers this effect, or whether an

Corresponding Author: Baoxia Zuo
School of Humanities, Tianjin University of Finance and Economics Pearl River College, P.R. China
©The Author(s) 2026. Published by BONI FUTURE DIGITAL PUBLISHING CO.,LIMITED. This is an open access article under the CC BY License(<https://creativecommons.org/licenses/by/4.0/>)

integrated model holds empirically. This study aims to fill these gaps by constructing and testing a "reference-point-task-framing-risk-aversion" integrated model grounded in Prospect Theory, thereby offering a cognitive-decision account and empirical support for intervening in college classroom oral silence.

2. Literature Review

2.1 Definition and classification of risk avoidance in foreign language classroom speaking

"Risk aversion," originating in behavioral economics, denotes the tendency to choose options that avoid potential losses over those that pursue equivalent gains under uncertainty (Kahneman & Tversky, 1979). In second language acquisition, this concept has been adopted to explain learners' conservative tendencies in classroom oral interaction. MacIntyre et al. (1998) noted that willingness to communicate is not a stable trait but is shaped by real-time assessment of the "risk of speaking." Accordingly, this study defines oral risk-avoidance as learners' psychological and behavioral tendency to refrain from speaking, simplify output, or remain silent, driven by anticipated negative outcomes such as public errors, face loss, or diminished evaluation.

This tendency manifests in two forms. Behavioral avoidance involves overt actions-lowering the head, avoiding eye contact, refraining from hand-raising, or abbreviating responses. Psychological avoidance entails internal withdrawal, such as using safe syntactic patterns, minimizing speech length when called upon, and avoiding personal opinions (Gałajda, 2017). These two forms may co-occur or appear independently. Importantly, risk-avoidance is not synonymous with low language proficiency; its prevalence even among high-proficiency learners' points to a cognitive-decision mechanism that operates independently of linguistic competence.

2.2 Perspective of risk avoidance behavior in foreign language classroom speaking: the introduction of prospect theory

Existing research on classroom oral risk-avoidance has followed two main explanatory paths. The affective trait path (Horwitz et al., 1986) attributes silence to stable dispositions such as fear of negative evaluation and communication anxiety, yet fails to account for within-learner variability across tasks and situations (Dewaele, 2019). The social context path (Jackson, 2024) emphasizes external shaping by classroom power relations, face culture, and peer pressure, but risks overstating external constraints and treating learners as passive respondents rather than active cognitive appraisers.

Both approaches share a common blind spot: they neglect the cognitive decision-making process

underlying the choice to speak or remain silent. Prospect Theory (Kahneman & Tversky, 1979) addresses this gap through two core features-reference-point dependence (outcomes perceived as gains or losses relative to subjective benchmarks) and loss aversion (loss avoidance outweighs equivalent gain pursuit). In classroom terms, learners confronting oral tasks do not passively "feel anxious and then fall silent"; rather, they undergo rapid cognitive appraisal-activating subjective reference points (e.g., "I should perform as well as peers," "the teacher expects a correct answer"), evaluating the potential consequences of speaking, and triggering risk-avoidance if the outcome is perceived as falling into the "loss domain."

Recent studies have begun applying Prospect Theory to L2 contexts. Zhang and Rahimi (2023) demonstrated that loss-framed instructions ("avoid deduction") reduce participation more than gain-framed ones ("earn extra points"); Papi et al. (2019) identified "ideal L2 self" and "ought L2 self" as distinct reference-point types with divergent behavioral effects. However, most existing work remains confined to single-construct validation or experimental framing manipulations, leaving unanswered in the Chinese college classroom context: What reference points do learners actually draw upon? How do these reference points jointly shape risk-aversion? And what role does teacher task framing play in this process?

2.3 Factors influencing risk-averse behavior in foreign language classroom speaking: reference points and task frameworks

Drawing on Prospect Theory and L2 literature, the determinants of classroom oral risk-avoidance can be categorized into two sets: reference-point dependence and teacher task framing.

Reference-point dependence-the subjective benchmarks against which individuals evaluate outcomes-manifests in classroom settings through four sources: social comparison (peer performance; falling short is perceived as loss; Elliot & Thrash, 2002), self-expectation (personal goals; Liu, 2022), cultural norms (collective harmony and face concerns; public errors as social loss; Jackson, 2024), and teacher setting (evaluation criteria and task requirements; Li, 2024). Among these, teacher-set standards exert the most direct influence due to the authority of instructor evaluation.

Teacher task framing stems from the "framing effect" of Prospect Theory, whereby alternative linguistic descriptions of the same task alter decision-makers' reference points and risk preferences. Oral task presentations may adopt a gain frame (e.g., "This helps improve fluency") or a loss frame (e.g., "Poor preparation affects scores").

Empirical evidence indicates that gain frames elicit higher participation willingness and lower risk perception (Zhang & Rahimi, 2023). However, most existing studies treat task framing as an independent external variable, leaving unexplored whether it operates independently or indirectly by modulating reference-point activation. This question constitutes a central impetus for the present study.

3. Research Questions

3.1 Inadequacies of previous studies

Existing research suffers from three major gaps. First, at the theoretical integration level, reference-point dependence and framing effects have been largely examined in isolation; no study has systematically incorporated multiple reference-point sources and task framing into a single model within the Chinese college classroom context. Second, at the mechanism level, the transmission pathway from reference points to risk-aversion remains unclear—specifically, whether teacher task framing functions as a mediator, a moderator, or both—a distinction that bears directly on pedagogical intervention strategies. Third, at the boundary-condition level, learners' oral proficiency

likely moderates key relationships, yet no empirical work has tested whether high- and low-proficiency learners differ in reference-point sensitivity or framing responsiveness, leaving pedagogical implications underspecified. The present study constructs and tests an integrated model to address these deficiencies.

3.2 Research questions and theoretical models

Based on prospect theory and in view of the above research deficiencies, this study proposes the following two core research questions:

Research Question 1: Do foreign language learners' multiple sources of reference (peer comparison, self-expectation, cultural norms, and teacher setting) influence their risk-averse behavior tendencies in classroom speaking through the mediation of perceived teacher task framing? In other words, is the "reference point-task framework-risk aversion" integration model proposed in this study (as shown in Figure 1) valid?

Research Question 2: Does the learner's foreign language (oral) proficiency moderate the critical path in the above integration model? That is, does the core effect of the model vary with learners' foreign language proficiency?

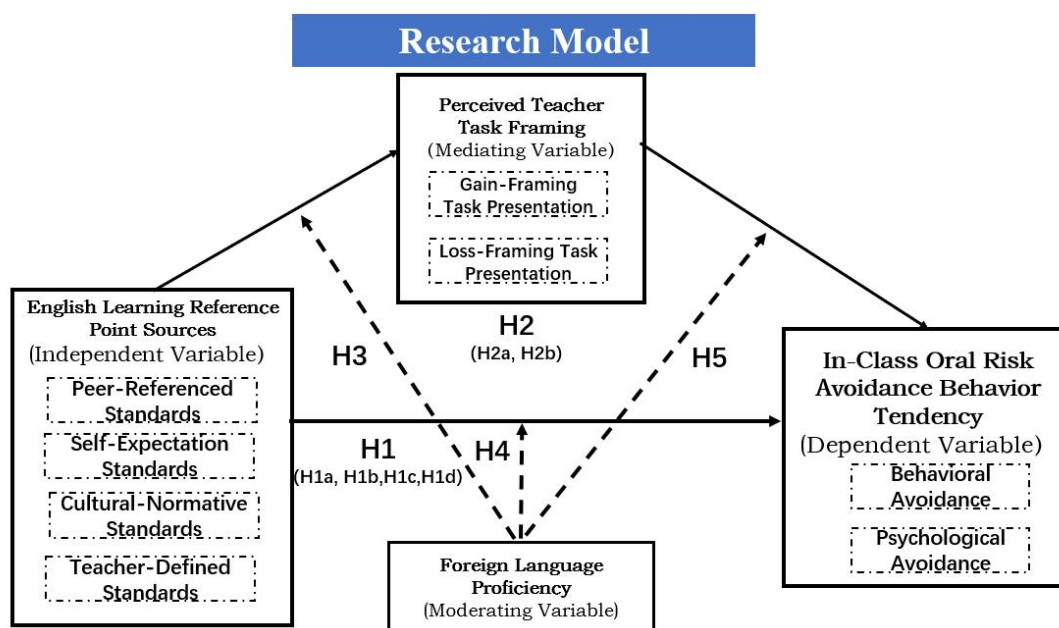


Figure 1. Figure of the Research Model (Integrated Model of "Reference Point-Task Framework-Foreign Language Proficiency-Risk Aversion")

Figure 1 Caption: Integrated Model of "Reference Point-Task Framework-Foreign Language Proficiency-Risk Aversion" for Classroom Oral English Participation.

Figure 1 Alt Text: A structural equation model diagram with six rectangular predictor boxes on the left pointing to two central oval constructs, which in turn point to two outcome oval constructs on the right, with standardized path coefficients shown on each arrow.

This figure presents the integration model proposed in this study. The figure contains four exogenous latent variables (peer comparison reference point, self-expectation reference point, cultural norm reference point, and teacher-set reference point), a mediator variable (perceived teacher task framework), an endogenous latent variable (risk-averse behavior tendency in classroom speaking), and a moderator variable (foreign language proficiency).

4. Method

4.1 Participants

A cross-sectional questionnaire design was used in this study, with Chinese undergraduate non-foreign language majors as the subjects. Through stratified random sampling, students from 11 universities in eastern, central and western China (including "double first-class initiative" universities, ordinary undergraduate and higher vocational colleges) were selected to participate in the survey. A total of 550 questionnaires were sent out and 541 valid questionnaires were recovered, with an effective recovery rate of 98.36%. The distribution of gender, grade, major, educational level and foreign language proficiency of the samples is shown in Table 1.

Table 1. Distribution of Demographic Characteristics of the Study Sample (N = 541)

Variable	Item	Frequency	Percentage (%)
Gender	male	194	35.9
	female	347	64.1
Educational Level	Vocational College student	88	16.3
	Undergraduate	453	83.7
	Humanities and Social Science	115	21.3
	Science	67	12.4
Major	Finance and Economics	107	19.8
	Engineering	251	46.4
	Arts	1	.2
	Sports	0	0
Grade	Freshman	307	56.7
	Sophomore	134	24.8
	Junior	77	14.2
	Senior	23	4.3
Foreign language proficiency	Pass CET4	380	70.2
	Not pass CET4	161	29.8
Total		541	100

4.2 Research tools

Three self-designed scales were employed, all using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Detailed psychometric properties are summarized below and in Table 2.

Foreign Language Learning Reference Point Source Scale. This 16-item scale comprises four dimensions: peer comparison, self-expectation,

cultural norms, and teacher setting (4 items each). In the pre-survey (n = 80) and formal sample (N = 541), Cronbach's α was .933. Confirmatory factor analysis (CFA) yielded $\chi^2/df = 2.657$, CFI = .911, TLI = .982, RMSEA = .140, SRMR = .071; AVE = .518, CR = .942, indicating good convergent validity.

Classroom Oral Risk Avoidance Behavior Tendency Scale. This 7-item scale comprises two

dimensions: behavioral avoidance (4 items) and psychological avoidance (3 items). Cronbach's α was .917. CFA results: $\chi^2/df = 2.176$, CFI = .977, TLI = .966, RMSEA = .068, SRMR = .026; AVE = .627, CR = .920.

Perceived Teacher Task Framework Scale. This 8-item scale comprises two dimensions: gain frame and loss frame (4 items each). Cronbach's α was .925. CFA results: $\chi^2/df = 2.065$, CFI = .910, TLI = .973, RMSEA = .157, SRMR = .073; AVE = .635, CR = .931.

Overall fit and construct validity. Across the three scales, χ^2/df values were below 3, CFI and TLI exceeded .90, and SRMR fell below .08, indicating acceptable model fit. The RMSEA values for the Reference Point Scale (.140) and Task Framework Scale (.157) exceeded conventional cutoffs (< .08 or < .10), likely due to scale complexity (estimating multiple parameters) and moderate sample size ($N =$

541), which tends to inflate RMSEA when chi-square is large (Hu & Bentler, 1999). Given that all other fit indices met recommended thresholds, factor loadings ranged from .62 to .86 ($p < .001$), and AVE and CR exceeded .50 and .70, respectively, construct validity was deemed acceptable. Such relaxation of RMSEA criteria for complex scales has precedent (Chen, 2007).

Common method bias and control variables. Harman's one-factor test on all scale items revealed that the first unrotated factor explained 28.6% of total variance, below the 40% threshold, suggesting no serious common method bias. Educational level (junior college vs. undergraduate), major (humanities/social sciences vs. non-humanities/social sciences), and grade (freshman to senior) were included as dummy-coded control variables in the analysis.

Table 2. Summary of Reliability, Validity and Confirmatory Factor Analysis Fit Index of Each Study Scale ($N = 541$)

Item	Statistics	Foreign Language Learning Reference Point Source Scale	Classroom Oral Risk Avoidance Behavior Tendency Scale	Perceived Teacher Task Frame Scale	Reference value of judgment standard
Reliability	Cronbach's α	.933	.917	.925	> .7 (Good)
Construct validity	KMO	.947	.924	.921	> .7
	Bartlett sphericity test (approximate chi-square)	6446.352***	2592.128***	3552.855***	$p < .001$
Convergent validity	AVE	.518	.627	.635	> .5
	CR	.942	.920	.931	> .7
CFA	χ^2/df	2.657	2.176	2.065	< 3
	CFI	.911	.977	.910	> .9
	TLI	.982	.966	.973	> .9
	RMSEA [90% CI]	.140 [.140, .154]	.068 [.068, .108]	.157 [.157, .189]	< .08 (Acceptable)
	SRMR	.071	.026	.073	< .08

Note: *** $p < .001$. Refer to Hu & Bentler (1999) for judgment criteria. Standardized factor loadings for all items on each scale ranged from 0.62 to 0.86 ($p < .001$). In order to control the length, the detailed item content and load matrix can be obtained from the corresponding author.

4.3 Data collection

Data collection took place from March to May 2026. First of all, the researcher explained the purpose and procedure of the study through the contact classes of foreign language teachers in various universities. With the consent of the teacher, the researcher (or a trained teaching assistant) used 10-15 minutes to organize the questionnaire before the end of the class. Before the start of the survey, the informed consent form was read out to the students, clearly informing them that the survey was anonymous and voluntary, and that all data were only

used for academic research and had nothing to do with course performance. Students have the right to withdraw at any time without any influence. After the questionnaire is completed, it will be collected on the spot. The entire data collection process followed the ethical principles of the Declaration of Helsinki and had passed the ethical review of the investigator's institution.

4.4 Data analysis

In this study, SPSS 27.0 and Amos 28.0 software were used for data analysis. The analysis process was divided into the following steps

according to the research questions:

4.4.1 Corresponding research question 1 (integrated model inspection):

Firstly, the data were preprocessed, including reverse item scoring conversion, continuous variable centralization and dummy variable coding. Secondly, descriptive statistics and correlation analysis were carried out to test the basic relationship between the core variables. Thirdly, structural equation modeling (SEM) was used to test the theoretical model presented in Figure 1, focusing on the significance of model fit index, path coefficient and mediating effect (using bias-corrected percentile Bootstrap method, repeated sampling 5000 times, calculating 95% confidence interval of indirect effect). Hu & Bentler (1999): $\chi^2/DF < 3$, CFI $> .90$, TLI $> .90$, RMSEA $< .08$, SRMR $< .08$.

4.4.2 Corresponding research question 2 (test of the moderating effect of foreign language proficiency):

To test the moderating role of foreign language proficiency in the theoretical model, CET-4 passing status (pass/fail) was employed as the operational indicator. CET-4 is a nationally recognized standardized English proficiency test in Chinese universities, with established relevance to classroom oral competence. The sample comprised 380 passers (70.2%) and 161 non-passers (29.8%).

Multi-group SEM was conducted in four steps. First, the sample was divided into passing and failing

groups. Second, measurement invariance was tested across three levels-configural, metric, and scalar-to ensure cross-group comparability of latent constructs. Third, path coefficients were compared between groups via chi-square difference tests; a significant $\Delta\chi^2$ ($p < .05$) indicates moderation. Fourth, bias-corrected percentile Bootstrap (5,000 resamples) was applied to compute 95% CIs for between-group coefficient differences; moderation is confirmed if the CI excludes zero. Model fit criteria followed Hu & Bentler (1999): $\chi^2/df < 3$, CFI $> .90$, TLI $> .90$, RMSEA $< .08$, SRMR $< .08$. All analyses were performed using SPSS 27.0 and Amos 28.0, with multi-group comparison procedures following Kline (2015).

5. Findings and Discussion

5.1 Structural equation model analysis

Structural equation modeling (SEM) was used to test the theoretical model presented in Figure 1. The results of model fitting test showed that $\chi^2/DF = 2.680$, CFI $= .950$, TLI $= .938$, RMSEA $= .083$, and SRMR $= .068$. Except for RMSEA, which is slightly higher than strict standard of .08, the other indicators reach a good fit level (Hu & Bentler, 1999). According to the comprehensive judgment, the fit between the model and the data is within the acceptable range, which supports further path analysis and mediation effect test.

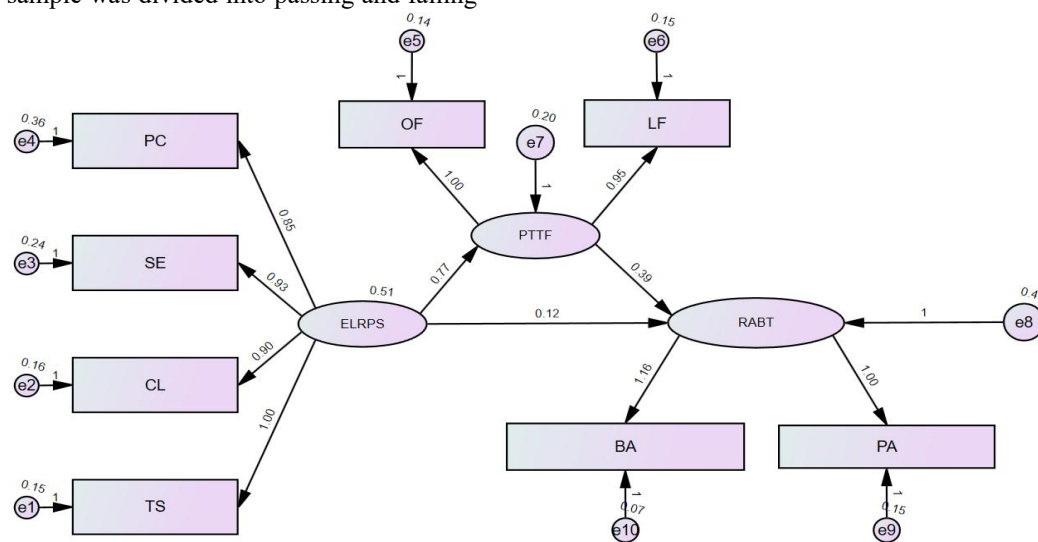


Figure 2. Path Coefficient of Structural Equation Model

Note: ELRPS: English Learning Reference Point Source; PTF: Perceived Teacher Task Frame; RABT: Risk Avoidance Behavior Tendency of Classroom Oral English; PC: Peer comparison; SE: Self-expectation; CL: Cultural norms; TS: Teacher setting; OF: Obtain Framing; LF: Loss Framing; BA: Behavior Avoidance; PA: Psychological Avoidance

Figure 2 Caption: Path Coefficients of the Structural Equation Model with Standardized Estimates.

Figure 2 Alt Text: A path diagram showing six latent variables connected by arrows with standardized regression weights ranging from negative 0.31 to 0.52, with the strongest positive paths from Peer Comparison and Self-Expectation to Perceived Teacher Task Frame, and from Obtain Framing to Risk Avoidance Behavior Tendency.

5.1.1 Direct effects of exogenous variables on endogenous variables

The results of the direct effect test of the four

reference point variables on the risk-averse behavior tendency of classroom speaking are shown in Table 3.

Table 3. Path Coefficient of Structural Equation Model and Hypothesis Test Results

Main effect path	β	p	Hypothesis	Conclusion
Main effect (H1)				
Peer comparison→Risk Avoidance Tendency	.708	***	H1a	Accept
Self-expectation→Risk Avoidance Tendency	.804	***	H1b	Accept
Cultural norms→Risk Avoidance Tendency	.850	***	H1c	Accept
Teacher setting→Risk Avoidance Tendency	.876	***	H1d	Accept
Mediating effect (H2)				
Total effect (c)	.448	***	-	-
Direct effect (c')	.203	***	-	-
Indirect effects (a*b)	.245	***	H2	Partial Mediating effect
(Proportion of mediating effect)	(54.70%)			
Reference Point Source×Perceived Teacher Task Obtain Frame				
→Risk Avoidance Tendency				
Reference Point Source×Perceived Teacher Task Loss Frame	.882	***	H2a	Accept
→Risk Avoidance Tendency				
Reference Point Source×Perceived Teacher Task Obtain Frame	.865	***	H2b	Accept
→Risk Avoidance Tendency				

Note: ** $p < .001$ *

From the perspective of effect size, the four types of reference points all positively predict learners' risk aversion tendency of classroom speaking, and the strength of prediction is different. The influence of the reference point set by the teacher was the most prominent ($\beta = .876$), followed by the reference point of cultural norms ($\beta = .850$) and the reference point of self-expectations ($\beta = .804$), and the influence of the reference point of peer comparison was relatively small but still significant ($\beta = .708$). This result shows that in the context of Chinese college foreign language classroom, teachers'

evaluation criteria and task requirements constitute the most dependent reference criteria for learners to judge "opening risk", and the influence of teachers' authority on students' decision-making in teaching interaction cannot be ignored.

5.1.2 Indirect effects of mediating variables

To examine the mediating role of the perceived teacher task framework in the reference point dependence → risk aversion" relationship, the Bootstrap method (repeated sampling 5000 times) was used to calculate the 95% confidence interval of the indirect effect. The results are shown in Table 4.

Table 4. Analysis of the Mediating Effect of Perceived Teacher Task Framing

Effect type	Effect value	95% CI	SE	p	Proportion of effect
Total effect (c)	.448	[.354, .541]	.047	***	-
Direct effect (c')	.203	[.081, .325]	.062	***	45.3%
Indirect effects (a*b)	.245	[.126, .279]	.039	***	54.7%

Note: ** $p < .001$ *

The indirect effect was significant (95% CI excluded 0), with the mediating effect accounting for 54.7% of the total effect, indicating partial mediation. This suggests that over half of the influence of reference-point dependence on oral risk-aversion

operates through learners' perception of teacher task framing. While internal reference points matter, how teachers verbally frame tasks-and where they set the gain/loss boundary-critically determines whether and with what intensity these reference points translate

into actual avoidance behavior.

Interaction analyses further revealed that both gain and loss frames significantly amplified the effect of reference-point dependence on risk-aversion (reference point $\times$ gain frame: $\beta = .882$, $p < .001$; reference point $\times$ loss frame: $\beta = .865$, $p < .001$), forming a "cognitive-situational reinforcement loop": internal reference points establish the baseline risk perception, while task framing serves as a contextual catalyst that magnifies this cognitive assessment into behavioral outcomes. Notably, the two framing types yielded similar effect sizes, suggesting that even gain frames-generally viewed as positive-may fail to alleviate risk-aversion among strongly reference-point-dependent learners; instead, they may intensify the psychological burden by reinforcing comparisons between task performance and reference standards.

5.2 The moderating effect of foreign language proficiency

In order to test the moderating effect of foreign language proficiency in the theoretical model shown in Figure 1, this study takes the CET-4 passing status of learners as the operational index of foreign language proficiency, and divides the sample into the passing group ($n = 380$, accounting for 70.2%) and the failing group ($n = 161$, accounting for 29.8%). Multi-group structural equation modeling was used to compare paths across groups.

5.2.1 Measurement invariance test

Before comparing the cross-group path coefficients, we first test whether the measurement model is invariant between the two groups to ensure that the reflection of the observed variables on the latent variables is comparable between the different groups. The test results are shown in Table 5.

Table 5. Measurement Invariance Test Results

Model	χ^2	df	χ^2/df	CFI	RMSEA	SRMR	Model comparison	$\Delta\chi^2$	Δdf	p
Model 1: The form remains unchanged	1024.631	476	2.152	.933	.063	.059	—	—	—	—
Model 2: Factor load invariance	1051.204	496	2.119	.931	.062	.063	Model 2 vs. Model 1	26. 573	20	.148
Model 3: Intercept invariance	1080.315	516	2.094	.929	.061	.064	Model 3 vs. Model 2	29. 111	20	.085

Note: $\Delta\chi^2$ is the chi-square value and Δdf is the degree of freedom difference.

As shown in Table 5, from Model 1 to Model 3, the CFI is higher than .90, the RMSEA is lower than .08, and the SRMR is lower than .08, indicating that the invariance models at all levels fit the data well. There was no significant difference in chi-square between model 2 and model 1 ($\Delta\chi^2 = 26.573$, $\Delta df = 20$, $p = .148$), indicating that factor loading was invariant between the two groups (metric invariance was established). There was no significant difference in chi-square between model 3 and model 2 ($\Delta\chi^2 = 29.111$, $\Delta df = 20$, $p = .085$), indicating that the intercept was invariant between the two groups (scalar invariance). Measurement invariance is established, which provides a

methodological premise for further cross-group path coefficient comparison.

5.2.2 Multi-group structural equation modeling path comparison

On the basis of measurement invariance, a multi-group structural equation model was constructed to estimate the standardized coefficients of each theoretical path in the passing group and the failing group, respectively. The overall fit of the model is good: $\chi^2/DF = 2.137$, $CFI = .928$, $TLI = .920$, $RMSEA = .064$, and $SRMR = .062$. The test results of the path coefficient and the difference between the two groups are shown in Table 6.

Table 6. Path Coefficient and Cross-Group Difference Test of Two Groups of Models

Path	Group of pass (n=380)	Group of not pass (n=161)	$\Delta\chi^2$ (1)	p	Is the regulatory effect significant?		
	β	p	β	p			
Peer comparison→Risk Avoidance Tendency	.645	***	.812	***	5.168	.023	Yes
Self-expectation→Risk Avoidance Tendency	.720	***	.892	***	6.203	.013	Yes
Cultural norms→Risk Avoidance Tendency	.788	***	.905	***	4.012	.045	Yes
Teacher setting→Risk Avoidance Tendency	.831	***	.903	***	2.891	.089	Not
Reference Point → Task Obtain Frame	.713	***	.758	***	1.547	.214	Not
Task Obtain Frame →Risk Avoidance	.524	***	.406	***	4.835	.028	Yes

Note: * $p < .001$; $\Delta\chi^2 (1)$ is the chi-square value (degrees of freedom = 1) of the constraint path coefficient compared to the free model when it is equal between the two groups.

Multi-group comparison revealed significant between-group differences in four of the six theoretical paths.

First, the three direct "reference point → risk-aversion" paths were all moderated by proficiency. The peer comparison ($\Delta\chi^2 = 5.168$, $p = .023$), self-expectation ($\Delta\chi^2 = 6.203$, $p = .013$), and cultural norms ($\Delta\chi^2 = 4.012$, $p = .045$) paths all showed larger coefficients in the failing group than in the passing group (peer: $.812 > .645$; self: $.892 > .720$; cultural: $.905 > .788$). This pattern indicates that lower-proficiency learners rely more heavily on reference points in their speaking decisions-attending more closely to peer performance, self-expectations, and face concerns-with weaker language ability itself constituting an additional source of decision-making risk that amplifies reference-point weighting.

Second, the "task framing → risk-aversion" path was also significantly moderated ($\Delta\chi^2 = 4.835$, $p = .028$), with the passing group showing a stronger coefficient ($\beta = .524$) than the failing group ($\beta = .406$). Higher-proficiency learners, possessing greater cognitive resources, appear more sensitive to the semantic nuances of task framing, whereas for

lower-proficiency learners, the inherent difficulty of language production already constitutes the primary "loss" source, leaving less room for framing to exert additional regulation.

Third, two paths showed no significant group differences: "teacher setting → risk-aversion" ($\Delta\chi^2 = 2.891$, $p = .089$) and "reference point → task framing" ($\Delta\chi^2 = 1.547$, $p = .214$). Notably, the teacher-setting path exhibited the strongest effects in both groups ($\beta = .831$ for passers; $\beta = .903$ for non-passers), suggesting that teacher-set evaluation criteria function as a foundational reference point that transcends proficiency levels-a universal product of classroom power relations and evaluation systems rather than a proficiency-dependent mechanism.

5.2.3 Cross-group comparison of mediating effects

In order to further investigate whether foreign language proficiency moderates the overall effect strength of the mediation path (reference point dependence → task frame → risk aversion), the indirect effect, direct effect and total effect of the two groups were calculated respectively, and the significance of the difference between the two groups was tested by Bootstrap method (repeated sampling 5000 times). The results are shown in Table 7.

Table 7. Decomposition and Cross-group Comparison of Mediating Effects in Two Groups

Type of effect	Group of pass (n=380)	Group of not pass (n=161)	Differences between groups			
	Effect value	95% CI	Effect value	95% CI	Difference	95% CI (Difference)
Total effect (c)	.374	.251, .512 .25, .512	.307	.182, .439 .182, .439	.067	-.041, .178 -.041, .178
Direct effect (c')	.312	.176, .448 .176, .448	.468	.294, .639 .294, .639	-.156	-.312, -.008 -.008, -.312
Indirect effects (a*b)	.686	.549, .823 .549, .823	.775	.612, .932 .612, .932	-.089	-.241, .063 -.241, .063

As shown in Table 7, the between-group difference for the indirect effect was not significant (95% CI included 0), whereas that for the direct effect was significant (95% CI excluded 0). This indicates that the mediating function of task framing operates consistently across proficiency levels, equally effective in translating reference-point dependence into avoidance behavior. However, the direct "reference point → risk-aversion" effect was significantly stronger in the failing group ($c' = .468$) than in the passing group ($c' = .312$), suggesting that for lower-proficiency learners, reference-point dependence triggers avoidance more directly, with less cognitive mediation by task framing. When language ability is insufficient, the reference-point impulse bypasses or compresses the processing chain; when proficiency is higher, the task framework has greater room to regulate this process.

Together, these findings confirm that foreign language proficiency exerts a path-specific moderating role: lower-proficiency learners show heightened sensitivity on the direct reference-point–risk-aversion pathway, while higher-proficiency learners show stronger responsiveness on the task-framing–risk-aversion pathway. This refined pattern provides a more nuanced empirical basis for understanding how proficiency shapes classroom speaking decisions.

5.2.4 Result discussion and mechanism explanation

Multi-group SEM yields three theoretical insights.

First, lower-proficiency learners show heightened reference-point sensitivity. The three reference-point→risk-aversion paths were consistently stronger in the CET-4 failing group, aligning with Hiver & Al-Hoorie (2020):

low-proficiency learners' cognitive resources are disproportionately occupied by risk monitoring, causing reference-point information to be processed as loss signals that amplify avoidance.

Second, higher-proficiency learners show greater framing responsiveness. The task-framing→risk-aversion path was stronger among passers ($\beta = .524$ vs. $.406$). Richer cognitive resources enable finer detection of gain/loss distinctions, echoing Zhang & Rahimi (2023) that proficiency underpins individual differences in framing effects.

Third, the "teacher setting" effect demonstrates cross-level stability. This was the only path showing no between-group difference yet exerted the strongest effects in both groups ($\beta > .83$). Teacher criteria function as a foundational reference transcending proficiency—a universal product of classroom power relations, implying that optimizing assessment language benefits all learners.

In sum, proficiency exerts path-specific moderation, grounding the dynamic "cognition–situation–ability" interplay.

5.3. Discussion

Four core findings emerge from this study: (1) all four reference-point types significantly and positively predict oral risk-aversion, with teacher setting exerting the strongest effect ($\beta = 0.876$); (2) perceived teacher task framing partially mediates the reference-point–risk-aversion relationship, accounting for 54.7% of the total effect; (3) task framing and reference-point dependence interact significantly, forming a "cognitive-situational reinforcement loop"; and (4) foreign language proficiency exerts path-specific moderation on key model pathways. These are discussed in turn below.

5.3.1 The main effect of reference point dependence: the dual prominence of teacher authority and cultural norms

Teacher-setting reference points exerted the strongest effect on risk-aversion ($\beta = 0.876$), consistent with Li's (2024) observation that teacher standards are internalized as the primary frame of reference in Chinese college EFL classrooms. Cultural norms also showed a robust effect ($\beta = 0.850$), confirming Jackson's (2024) claim that face culture renders public errors a social loss to be avoided. Notably, the teacher-setting path showed no significant difference between proficiency groups (see Table 6), indicating that teacher authority operates universally across proficiency levels.

5.3.2 The mediating role of task framework: from cognitive assessment to behavioral output

The mediating effect of perceived teacher task framing (54.7%) surpassed Zhang & Rahimi's (2023) experimentally observed direct effect of task framing. While their work demonstrated the direct impact of gain/loss framing on behavior, it did not clarify how framing interacts with learners' preexisting reference points. Our mediation model shows that task framing does not operate independently of cognitive structure but serves as a "transmitter" linking reference-point dependence to behavioral output. This aligns with Dörnyei's (2009) L2 motivational self-system premise that forward-looking reference points require situational experience to translate into actual behavioral tendencies.

5.3.3 The "cognitive-situational reinforcement circuit": A theoretical extension to the framing effect

Both gain and loss frames significantly amplified the effect of reference-point dependence on risk-aversion (interaction $\beta \approx 0.87$), revealing a "cognitive-situational reinforcement loop" previously unexamined in the literature. This finding engages Kahneman & Tversky's (1979) original framing theory: while framing effects show predictable directionality in "clean" decision contexts without reference-point dependence, classroom settings involve learners entering tasks with preexisting reference points, which substantially moderates the direction and intensity of framing effects. Gain frames—ostensibly positive—may paradoxically activate rather than alleviate risk-aversion among strongly reference-point-dependent learners by implicitly prompting comparisons between current performance and a higher "ideal state." This challenges the simplistic assumption that gain frames are uniformly beneficial.

5.3.4 Pathway-specific regulation of foreign language proficiency: three different sensitive modes

Multi-group SEM revealed path-specific, rather

than global, moderation by proficiency.

Mode 1: Lower-proficiency learners' heightened reference-point sensitivity. The peer comparison, self-expectation, and cultural norms→risk-aversion paths were significantly stronger in the CET-4 failing group. Language deficits impose higher cognitive load, causing reference-point information to be processed as loss signals that amplify avoidance.

Mode 2: Higher-proficiency learners' enhanced framing responsiveness. The task-framing→risk-aversion path was stronger among passers ($\beta = 0.524 > 0.406$), as richer cognitive resources enable finer discrimination of framing nuances.

Mode 3: Cross-level stability of teacher setting. This path showed no between-group difference yet exerted the strongest effects in both groups ($\beta > 0.83$), positioning teacher-set standards as a foundational reference transcending proficiency.

These findings complement Horwitz et al.'s (1986) anxiety account: anxiety is not the "first cause" of silence but the emotional product of reference-point–task-frame interaction. Learners first appraise "speaking may incur loss," which triggers anxiety and reinforces avoidance (Dewaele, 2019). In sum, reference-point dependence is the core cognitive antecedent, task framing mediates and reinforces, and proficiency regulates through three sensitivity modes—offering an integrative framework for understanding classroom interaction quality.

6. Conclusion

6.1 Main findings

This study constructs and tests an integrated "reference-point–task-framing–proficiency" model grounded in Prospect Theory. Four main findings emerge.

First, reference-point dependence serves as the core cognitive antecedent of oral risk-aversion. All four reference-point types positively predicted risk-aversion, with teacher setting exerting the strongest effect ($\beta = 0.876$), highlighting the primacy of teacher evaluation criteria in students' risk perception.

Second, perceived teacher task framing partially mediates this relationship (54.7%), functioning as the key link that transmits internal reference points into behavioral output through the teacher's gain/loss framing of tasks.

Third, task framing and reference-point dependence interact significantly, forming a "cognitive-situational reinforcement loop." Both gain and loss frames amplify reference-point effects, challenging the assumption that gain frames are uniformly beneficial.

Fourth, proficiency exerts path-specific moderation: non-passers show stronger direct

reference-point→risk-aversion paths (heightened reference-point sensitivity), passers show a stronger task-framing→risk-aversion path (enhanced framing responsiveness), while the teacher-setting→risk-aversion path remains stable across proficiency levels.

6.2 Teaching inspiration

First, use hierarchical task framing. Gain frames may backfire among reference-point-dependent learners; neutral frames ("no right or wrong") are preferable for non-passers, while gain frames suit passers, supplemented by scaffolding for both.

Second, restructure reference-point reliance. Explicit activities can help students shift from social comparison to self-comparison and process-based standards. Optimizing teacher assessment language remains a priority due to its cross-level influence.

Third, build a low-risk, high-support community. Teachers should model risk-taking, position the classroom as a "learning laboratory," and provide tailored support, especially for non-passers.

6.3 Research limitations and future prospects

This study has four limitations. First, the cross-sectional design precludes causal inference; future research should employ longitudinal or experimental designs. Second, CET-4 status primarily reflects reading and listening comprehension, offering limited validity for oral proficiency; subsequent studies should incorporate standardized oral test scores or multidimensional proficiency measures. Third, self-report data are susceptible to social desirability bias; triangulation with behavioral observations and teacher evaluations is recommended. Fourth, the static measurement of reference points fails to capture their dynamic fluctuations; experience sampling methods could track real-time reference-point shifts during specific tasks. These limitations inform the above future research directions.

Conflict of Interest

The author declares that she has no conflicts of interest in this work.

Acknowledgement

This Research was Supported by China Foreign Language Education Fund (Project No:ZGWYJYJJ13A156,"A Study on Classroom Risk Decision-Making Mechanisms of English Learners and Teaching Interventions Based on Prospect Theory").

References

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.

Dewaele, J.-M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among foreign language learners. *Journal of Language and Social Psychology*, 38(4), 523–535.

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, 82(5), 804–818.

Galajda, D. (2017). *Communicative behavior of a language learner: Exploring willingness to communicate*. Springer.

Hiver, P., & Al-Hoorie, A. H. (2020). *Research methods for complexity theory in applied linguistics*. Multilingual Matters.

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

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.

Jackson, J. (2024). *Language and intercultural communication in the new era*. Routledge.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Li, C. (2024). Teacher feedback and learner perception in the Chinese EFL classroom: A dynamic interaction. *Language Teaching Research*, 28(1), 45–63.

Liu, H. G. (2022). A study on reference points of Chinese college students' engagement in English learning: Types, influences and interventions. *Foreign Language Circles*, (3), 45–52.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in an L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562.

Papi, M., Bondarenko, A. V., Mansouri, S., Feng, L., & Jiang, C. (2019). Rethinking L2 motivation research: The 2×2 model of L2 self-guides. *Studies in Second Language Acquisition*, 41(2), 337–361.

Wang, D. (2021). A study on classroom silence of foreign language learners from the perspective

- of L2 motivational self system. *Foreign Language and Foreign Language Teaching*, (2), 56–64.
- Wen, Q. F. (2014). The output-driven and input-contributing hypothesis: An attempt to construct a college foreign language classroom teaching theory. *Foreign Language Circles*, (2), 1–12.
- Wen, Z. (2018). Empowerment: A new approach to foreign language classroom teaching reform. *Foreign Languages*, 34(3), 1–7.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage Publications.
- Xu, J. F., & Kou, J. N. (2017). Research on the relationship between college foreign language classroom learning environment and learners' willingness to communicate in oral English. *Foreign Language and Foreign Language Teaching*, (2), 39–49.
- Xu, J. F., & Nie, R. (2020). Hotspots, trends and prospects of foreign language education psychology research. *Frontiers of Foreign Language Education Research*, 3(3), 47–54.
- Zhang, L. J., & Rahimi, M. (2023). EFL learners' willingness to communicate in class: The role of task framing and perceived classroom climate. *System*, 112, 102956.

How to Cite: Zuo, B. (2026). Understanding Classroom Silence: A Cognitive-Decision Model of Reference-Point Dependence and Task Framing. *Contemporary Education and Teaching Research*, 07(09), 278-293.
<https://doi.org/10.61360/BoniCETR262020690903>

Appendix:**Risk Avoidance Behavior Scale for Foreign Language Learner**

(Foreign Language Learning Reference Point Sources Scale,
In-Class Oral Risk Avoidance Behavior Scale,
Perceived Teacher Task Framing Scale)

Dear students: Hello!

You are cordially invited to participate in a scientific research project. This test is mainly used to investigate the reference points of risk avoidance behavior of foreign language learners. This questionnaire is an anonymous survey. All data will be used for scientific research only and will be kept strictly confidential. Your truthful answers are crucial to this study. Thank you for your support and cooperation!

Please tick the circle after each question that matches your answer. Note that only one new answer can be selected for each question. There is no right or wrong answer. Please complete it independently and do not consult with others. When answering questions, please do not hesitate or think too long. For questions that are not easy to answer, please make a choice that is closer to the actual situation.

Thank you again for your support and cooperation.

In order to make the survey more valuable, you should fill in the following information:

Basic information:

Your gender:

- ☐ Male
☐ Female

Your current level of education:

- ☐ College (Higher Vocational)
☐ Undergraduate

Your major:

- ☐ Humanities and Social Sciences Major
☐ Natural Science Major
☐ Engineering majors
☐ Major in Finance and Economics
☐ Art Majors

☐ Sports majors

Name of your university:

Region of your university:

Grade you belong to:

- ☐ First year of university
☐ Sophomore year
☐ Junior year of university
☐ Senior year of university

CET-4 passing status

- ☐ Passed
☐ Not Passed

1 Foreign Language Learning Reference Point Sources Scale (FLL-RPSS)**Instructions**

Please indicate the extent to which each statement describes your typical experience in English learning.

1 = Strongly disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly agree

Peer-Referenced Standards

1. I usually judge how well I am learning English by comparing my performance with that of my classmates.
2. Other students' English performance serves as an important reference when I evaluate my own English

ability.

3. After learning about my classmates' English performance, I often reassess my own learning outcomes.
4. (*Reverse-coded*) When evaluating my English learning, I rarely use my classmates' performance as a reference.

Self-Expectation Standards

5. I have a clear internal standard for the level of English proficiency I believe I should achieve.
6. I usually evaluate my current English performance based on my personally expected level.
7. I assess my performance in English exams or classroom tasks according to the goals I set for myself in advance.
8. My past English learning experiences provide an important reference for evaluating my current performance.

Cultural-Normative Standards

9. In today's society, being proficient in English is widely regarded as a basic competence for university students.
10. Societal expectations regarding English proficiency influence how I evaluate my own English learning.
11. The surrounding social and academic environment leads me to view English achievement as an indicator of my competitiveness.
12. The social value attached to English proficiency serves as an important reference point in my self-evaluation.

Teacher-Defined Standards

13. Teachers' grading criteria or overall performance benchmarks are important references for evaluating my English performance.
14. Teachers' explicit requirements for "good performance" often serve as standards by which I judge my learning outcomes.
15. I evaluate my performance by taking into account teachers' explanations of task difficulty or assessment expectations.
16. Teachers' feedback and evaluative comments directly influence my overall judgment of my English performance.

2 In-Class Oral Risk Avoidance Behavior Scale (ORA)

Instructions

Please recall your actual behavior in English classes and indicate how often each statement applies to you.

1 = Never 2 = Rarely 3 = Sometimes 4 = Often 5 = Always

Behavioral Avoidance

1. Unless I am very confident about my answer, I usually avoid volunteering to speak in English in class.
2. During oral practice, I tend to use simple expressions with low risk of error rather than attempting unfamiliar or complex forms.
 4. In group discussions, if I feel that my ideas may be difficult to express clearly in English, I usually choose to remain silent.
 5. Due to concerns about possible negative evaluation of my pronunciation or grammar, I reduce my frequency of speaking English in class.

Psychological Avoidance

3. When required to give an oral presentation or perform a dialogue in front of the class, I tend to avoid the task or hope that the difficulty will be reduced.
 6. When facing a challenging oral task, my initial focus is often on the risk of making mistakes rather than the potential learning benefits.
 7. (*Reverse-coded*) In English class, I am willing to try expressing ideas that I have not yet fully mastered.
-

3 Perceived Teacher Task Framing Scale (PTTF)

Instructions

Based on your perception of how your English teacher typically presents classroom tasks and evaluations, please indicate your level of agreement with the following statements.

1 = Strongly disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly agree

Gain-Framed Task Presentation

1. My English teacher usually presents classroom tasks as opportunities to develop language ability and gain learning experience.
 2. When encouraging participation, the teacher emphasizes what students can learn from the task rather than the mistakes they might make.
 3. In feedback, the teacher typically highlights students' progress or positive aspects of their performance.
 4. When assigning tasks, the teacher emphasizes that the tasks allow for diverse or personalized forms of expression.
-

Loss-Framed Task Presentation

5. When explaining tasks or exams, the teacher often emphasizes where points may be lost or errors are likely to occur.
6. The teacher highlights task difficulty, average scores, or pass rates to caution students against performing poorly.
7. The teacher's feedback tends to focus first on students' shortcomings or gaps relative to established standards.
8. By comparing students with other classes or previous cohorts, the teacher emphasizes the potential disadvantages of weak performance.