Open Access Full Text Article

ORIGINAL RESEARCH

The Effects of General Listening Anxiety and Listening Test Anxiety on Self-Perceived Listening Performance Among Chinese English Learners

Ke Li [1,2], Xiaoqing Qin[3], Suhe Ji[3], Linfeng Zou[4]

[1]School of Journalism and Communication, Wuhan University, Wuhan, Hubei Province, People's Republic of China; [2]Center for Studies of Media Development, Wuhan University, Wuhan, Hubei Province, People's Republic of China; [3]School of Foreign Languages, Central China Normal University, Wuhan, Hubei Province, People's Republic of China; [4]School of Communication, Hankou University, Wuhan, Hubei Province, People's Republic of China

Correspondence: Suhe Ji, School of Foreign Languages, Central China Normal University, Wuhan, Hubei Province, 430079, People's Republic of China, Email jisuhe@mails.ccnu.edu.cn; Linfeng Zou, School of Communication, Hankou University, Wuhan, Hubei Province, 430212, People's Republic of China, Email cmzoulinfeng@hkxy.edu.cn

Purpose: This research seeks to conceptualize foreign language listening anxiety (FLLA) and provide its measurement, and further explore the influences of FLLA on self-perceived listening performance.

Methods: In Study 1, semi-structured interviews explored FLLA-arousal situations. Follow-up reliability and validity tests for the newly-developed scale were testified. In Study 2, structural equation modeling explored the relationship between FLLA and self-perceived performance, which was followed by the comparison of the effects of different types of FLLA on self-perceived performance between English and non-English major students.

Results: The componential factors of FLLA included two factors, namely general listening anxiety and listening test anxiety, and general listening anxiety was represented by FLLA in classroom, daily usage, and media learning. The results also showed that listening test anxiety negatively affected self-perceived performance; general listening anxiety positively affected listening test anxiety but did not affect self-perceived performance, and listening test anxiety played a full mediation role. Moreover, findings revealed that non-English major students' general listening anxiety was higher than that of English major students. However, the multi-group analysis showed that these two groups did not differ in the effect of general listening anxiety on listening test anxiety, and in the effects of listening anxieties on self-perceived performance. For the two groups, the mechanism of anxiety-and-performance relationship was consistent.

Conclusion: The results of this research have expanded the knowledge of listening anxiety by distinguishing general listening anxiety from listening test anxiety. Moreover, by testifying the mediator of listening test anxiety, this research deepened the understanding of the effects of different types of FLLA on self-perceived listening performance and the intensity differences of listening anxieties in English and non-English majors. Furthermore, the research has contributed to the literature on FLLA research based on Complex Dynamic Systems Theory, and has practical pedagogical implications for future studies.

Keywords: foreign language listening anxiety, general listening anxiety, listening test anxiety, structural equation modeling, multi-group analysis

Introduction

The relationship between anxiety and language achievement has attracted the attention of researchers for decades.[1] In language education, most of the foreign language anxiety (FLA) and foreign language listening anxiety (FLLA) research has focused on anxiety elicited in the classroom[1,2] and test situations.[3,4] The practical reason for this may attribute to the limited access to the target language outside the classroom.[5] But there is insufficient research on FLA and FLLA in out-of-school exposure situations. However, with the development of technology and media, language learners have more opportunities to access English outside of class. Research into out-of-school English-as-a-foreign-language (EFL) learning has found that young children learned English

outside the school due to the prominent role of English in the internet, audiovisual resources, and various media.[6] Research also found that the pervasiveness of English exposure outside the school makes English play a dominant role in daily life in European countries.[7] Maybe it is the influence of linguistic distance (European languages are typologically similar to the target language English) that contributes to the strong facilitative role of out-of-school exposure to EFL learning.[7] Although Chinese is typologically distant from English, frequent and extensive exposure to English due to the audiovisual input-rich situations nowadays may compensate for linguist distance. Frequent exposure to English outside the school may diminish FLA in classroom learning and test taking. Consequently, previous research lacks exploration of FLA and FLLA in extended situations, which may provide greater flexibility for further research to bridge the gap.

Previous research conducted from the theoretical perspective has mainly focused on exploring the conceptual defining of FLLA, measurement of FLLA, and the mechanism of FLLA-and-performance relationship. However, it was found that the defining of FLLA was unclear-cut and non-transparent, and the measurement of FLLA was inaccurate and inconsistent to the conceptual definition to a great extent.[8] Research into the mechanism of FLLA-and-performance has mainly examined the linear effect of FLLA on listening performance; however, findings concerning FLLA-and-performance are inconsistent. For example, a significant negative correlation was found between FLLA and self-perceived listening proficiency,[9] whereas it was also found that FLLA was not a predictor of listening performance.[10] Likewise, research found that listening test anxiety was negatively related to listening performance;[11] on the other hand, the non-significant result of the effect of listening test anxiety on listening performance was detected by previous research.[3] Against this backdrop of inconsistent findings, it is necessary to distinguish different types of FLLA because the conceptual confusion and contradictory of FLLA may introduce a major threat to the validity of the measure of FLLA; it is of great importance of developing a new scale to precisely assess FLLA according to the theoretical defining; also there is an urgent need for tapping into a unifying mechanism of FLLA-and-performance relationship based on exploring the complex relationship among different types of FLLA and listening performance.

As mentioned above, the present research aimed to explore FLLA from various situations, including in and outside the classroom, and the impacts of different types of FLLA on listening performance. We undertook two studies for the purposes of the FLLA dimension identification and measurement, and FLLA-and-performance relationship research. In Study 1, we firstly conducted semi-structured interviews to explore the potential anxiety-arousal situations and relationships between different types of FLLA, and further to develop a new scale for measuring FLLA from the situational perspective. In Study 2, using the newly developed scale, we verified the hypothesized relationships among different types of FLLA and self-perceived listening performance by employing structural equation modeling, and compared different effects of listening anxieties on listening performance among Chinese English major and non-English major learners. The main contributions of this research may lie in offering a new perspective to reveal the effect of FLLA on listening performance and deepening the understanding of the intensity differences of the impacts of different listening anxieties on listening performance.

Literature Review
Defining Foreign Language Listening Anxiety

The situation-specific perspective claimed that FLA was not a composite of other anxieties, but a distinct set of feelings of tension, apprehension, and nervousness limited to the language learning situations.[12,13] It can be seen as trait anxiety restricted to a given situation;[14,15] in other words, language learners had the trait of feeling state anxiety when learning and/or using language.[16] However, a criticism of situation-specific perspective was that the situation was not precisely and clearly defined.[14] Previous research has showed that language classroom and test situations were two situations that were especially prone to elicit language anxiety,[12,17] and FLLA evoked in classroom was considered general listening anxiety.[17]

Based on Complex Dynamic Systems Theory (CDST), general listening anxiety and listening test anxiety differed in the strength of attractor states.[18] General language anxiety was considered learners' self-perceived unpleasant attractor state in various situations[19] that was a rather stable tendency or outcome for the dynamic system.[18] While high-stakes test situations may contribute to the perturbations with the overall stable system, thereby changing the strong unpleasant attractor state into a weaker attractor state.[18] In addition, general listening anxiety and listening test anxiety differed in the effects on listening performance. General listening anxiety was elicited in various situations (such as in daily communication situations[5]) in which

listening comprehension was non-task-oriented and associated with anxiety-free responses.[20] It indicates that general listening anxiety may not have a direct effect on listening test performance. However, listening test anxiety directly and adversely affected task-oriented listening performance in high anxiety-eliciting test situations.[21,22] Therefore, the dynamic approach based on CDST[19] captures the dynamic characteristics of FLLA and reframes different types of FLLA. In this study, general listening anxiety refers to an unpleasant emotion associated with listening in a foreign/second language in various situations that tends to settle in over time; listening test anxiety refers to a feeling of tension in listening comprehension in tests that constantly fluctuates due to the importance of tests, high-stakes situations, and interactions between these factors.

Furthermore, the dynamic approach emphasizes a contextualized perspective to construct FLA. According to CDST, FLA was influenced by both learner factors and contextual effects,[16,18,19] and subtle or strong effects of situations can exert great impacts on FLA.[23] It was proposed to identify FLA in most typical situations in learners' past learning experiences because situational contexts have been confirmed to have unintended effects[18] and FLA may interact with the situation of which FLA was a part. Moreover, attention should be given to the evolution of language learning from classroom situation to situations on a daily basis[24] because with the development of technology and media learners currently have more opportunities and resources available on various media to learn a foreign language beyond academic settings.[25,26] In other words, general listening anxiety should extend from classroom to a dimension of anxiety aroused in a wide range of situations.

In sum, the situation-specific perspective on FLLA emerged as trait-focused characteristics at the beginning, and then the dynamic perspective transformed it into dynamic and context-dependent construction, foregrounding the importance of situations where language learning takes place and thereby distinguishing general listening anxiety from listening test anxiety. In this sense, CDST has provided evidence that identifying different types of FLLA can help to specify different levels of anxiety, leading to different effects on listening performance.

The Effects of Foreign Language Listening Anxiety on Listening Performance

Research over the past two decades has found that FLLA was negatively related to a variety of objective measures of foreign language listening achievements, such as listening proficiency tests,[17,27,28] course grades,[29,30] and self-designed listening comprehension tests.[10,31,32] However, results from the meta-analysis revealed that the strongest relation between FLA and language achievements was observed for self-perceived measures.[33,34] In FLLA research, it has been proved that FLLA and self-perceived listening proficiency were negatively related,[9] and FLLA was a powerful negative predictor of self-rated foreign language listening proficiency.[35] These findings suggest that FLLA may affect learners' self-perceived assessment of the performance gap between an ideal score and the actual obtained score, and this subjective test performance can be assessed by self-perceived test performance.

With regard to listening test anxiety, evidence showed that the intensity of listening test anxiety was significantly higher than general listening anxiety;[17] it indicated that language learners experienced a higher level of anxiety in high-stakes situations,[22] which echoes the assumption that the level of anxiety depends on the situation.[36] However, despite the significant negative relationship between test anxiety and achievement measures, non-significant results of some studies deserve more attention. It was found that test anxiety did not influence average academic achievement[37] or listening test performance.[3] These results may attribute to the delayed distribution time of listening anxiety scales. The long duration (eg, two weeks, a month, or a semester) between the listening anxiety tests and listening comprehension tests made it difficult to capture authentic feelings of listening test anxiety.[3,38] Instead, what a delayed distributed scale measured is the respondents' global self-perception of FLLA in various situations. It suggests that test anxiety should be measured right after the test, and the measurement of FLLA in different situations is of great significance because it can more effectively distinguish listening test anxiety from listening anxiety in other situations.

The Measures of Foreign Language Listening Anxiety

Currently the most widely adopted or modified scale is Kim's[9] Foreign Language Listening Anxiety Scale, followed by a scale developed by Elkhafaifi (2005).[30] However, the validity of these two scales was questionable due to the inconsistency between the definitions and measurements.[8] However, Horwitz, et al's[12] Foreign Language Classroom Anxiety Scale (FLCAS) and Chang's[17] Listening Anxiety Questionnaire are the scales that were developed based on

corresponding conceptual definitions. But FLCAS focused on overall language skills, serving as a macro scale to identify FLA; it seemed inappropriate to use FLCAS to measure FLLA. Chang's[17] scale was situation-specific-based in terms of both the conceptual defining and scale development. Two situations were identified, named the general situation and the listening test situation. Nevertheless, the general situation was not specified. Few words may shed a dim light on the explanation of the term of the general situation, which was equal to classroom. However, currently classroom may not represent all situations in which listening anxiety prevails. Therefore, to ensure the accuracy of the measurement, FLLA should be measured based on various situations.

When academic majors are taken into consideration, mixed results of measurements of FLLA are obtained from previous studies. Research has found that humanities and non-humanities majors did not differ on the level of FLLA.[9] On the other hand, significant differences were found between social science majors and math majors in FLLA, with math majors being more anxious than social science majors.[39] However, the measurement of different intensity of FLLA between English and non-English majors has been neglected. Previous studies have given great attention to the differences between English and non-English majors in motivation,[40,41] engagement,[42] and learning environment.[43] These differences may reveal different learning styles of EFL learners; thereby we believed that there may be some differences in FLLA between English and non-English major students.

Based on the above discussion, this research aimed to investigate the following research questions: (1) How is FLLA measured from the situational perspective? (2) How and to what extent does FLLA affect EFL learners' self-perceived listening performance? (3) How do Chinese EFL learners in different academic majors differ in FLLA? Two studies were developed for inquiry into the above three research questions.

Study 1
Research Objectives
To answer the first research question, Study 1 was conducted in the following three steps. The first step was to conduct semi-structured interviews that tapped into various anxiety-arousal situations to form an initial questionnaire. The second and third step involved an Exploratory Factor Analysis (EFA) and a Confirmatory Factor Analysis (CFA), respectively, which were conducted to testify reliability and validity of the newly developed FLLA scale.

Participants
Study 1 included three independent samples: the interviews sample, the EFA sample, and the CFA sample. In the semi-structured interview sample, 23 university students (12 English majors, 11 non-English majors) from grade one to grade four aged from 18 to 24 ($M = 21.15$, $SD = 1.713$) were recruited based on purposive sampling.[44] During the interview process, English majors and non-English majors were interviewed, respectively. It was found that no new theme can be generated after analyzing the interviews with 9 English majors; then three more interviews were conducted to achieve data saturation based on the guideline by Francis et al.[45] Likewise three more interviews were conducted after no additional issue was identified from 8 non-English majors' interviews. Among all the participants, 10 were males (43.48%) and 13 were females (56.52%); their listening proficiency levels differ from high-intermediate to low-intermediate. They had completed at least six years of English studies before entering the university. None of them had ever visited or lived in any English-speaking counties. The interview process revealed the universality of the anxiety-arousal situations encountered by English and non-English majors. In other words, both English and non-English majors identified similar general situations; they only differed in the degree of sensitivity of generating these anxiety-arousal situations, with English majors more sensitive to anxiety-arousal situations. In the EFA sample, 215 English majors from grades two and three in a university in a city in central China participated in this study. Among them, 22 were males (10.23%); 193 were females (89.77%); they ranged in age from 18 to 22 years ($M = 19.63$, $SD = 0.953$). Because English as a foreign language has been designated in the curriculum, English courses are compulsory from junior middle school to university (some primary schools begin to teach English as a foreign language from Grade three). Therefore, they had received at least 8–12 years of formal EFL education. In the CFA sample, 208 English majors aged from 18 to 22 years ($M = 19.80$, $SD = 0.844$) participated in the study. They had similar EFL learning backgrounds to that of the participants in the pilot EFA study. Among them, 12 were males (5.78%) and 196 were females (94.23%). Snowball sampling was

applied for the data collection process of CFA. The questionnaire was first distributed to English majors from two universities in central China, and then distributed to other English majors through invitations given by the initial respondents. The questionnaire was distributed in the form of network questionnaire through social media and group distribution. The quality control audit function was adopted in the questionnaire distribution; in other words, participants with missing data were automatically excluded.

Procedures

The semi-structured interviews adopted the situation-specific perspective on conceptualizing FLLA; more specifically, based on Chang's binary model,[17] FLLA was identified as general/classroom listening anxiety and listening test anxiety. FLLA in this study was defined as worry and apprehension of listening comprehension in the classroom, tests, and other out-of-school English exposure settings. To generate new items for developing FLLA scale, the structure of the interviews was developed based on previous research,[12,17] focusing on FLLA in the classroom, tests, and other potential situations. The interviews were conducted with individual participants, lasting 30–45 minutes each.

According to the participants' responses in the interviews, the initial 25 items were generated and then evaluated for appropriateness by three experts. The items with similar or ambiguous meanings were eliminated. After the initial piloting, the retained 18 items that consisted of four factors were used to measure FLLA. A Likert response format was adopted consisting of a 7-choice response scale with the highest number 7 corresponding to *strongly agree* and the lowest number 1 corresponding to *strongly disagree*.

Data Analysis

The data analysis consists of three major stages. First, according to Hesse-Biber's coding method, a top-down thematic analysis coding was employed to analyze the transcribed texts.[46] The analysis of the interview data progressed in three major steps:[46,47] (a) firstly each transcription was analyzed independently as raw materials; (b) subsequently patterns were identified by comparisons of the coding of raw materials; (c) then themes were generated from the coded patterns. Thus, an initial pool of scale items was formed based on the results of the interviews. Second, a total of 215 valid cases were coded and processed with SPSS,[48] and an EFA with the extraction method of principal axis factoring and rotation method of Oblimin with Kaiser Normalization was conducted to determine the component structure of the newly-developed scale. Third, a total of 208 cases were processed with AMOS[49] and a CFA was performed to evaluate construct validity and to testify whether the higher-order construct of general listening anxiety was reasonable.

Results

The Results of the Semi-Structured Interviews

Through the top-down template analysis process, four major thematic patterns were identified from the interview transcripts. The coding results of the semi-structured interview are presented in Table 1. Among them, two a priori themes detected from previous literature were *listening test anxiety* and *listening anxiety in classroom*, and two newly emerged patterns were *listening anxiety in daily usage* and *listening anxiety in media learning*. The results of interviews also revealed that the above anxiety-arousal situations are the representatives of general situations that may largely elicit FLLA. Although the participants differed in the degree of experiencing listening anxiety in these situations, they agreed that the above situations can provoke their FLLA if they were exposed to these situations.

Listening test anxiety referred to listeners' fear of failure in tests. *Listening anxiety in classroom* referred to EFL learners' feelings of uneasiness and nervousness in classroom performance. These results concurred with previous research[17,30] that classroom and test situations were commonly identified as the main anxiety-provoking situations. Participant 17 said that:

Sometimes I get worried before the English listening class because I was afraid to answer questions in English listening class.

Listening anxiety in daily usage referred to obstacles influencing successful receiving and decoding of information when using English on daily basis. Almost half of the participants reported that they were overly concerned about their reception of listening information when listening took place in daily communication, such as answering phone calls from

Table 1 Coding Results of the Semi-Structured Interviews

Key Themes	Sub-Themes	Transcript Samples
LTA	LTA	*"I am afraid of failing to keep up with the subsequent content when missing a key word in the listening test."* *"I always feel nervous about the high-stakes tests atmosphere, worry about unfamiliar topics, or not having enough time to think about what I've heard in the listening test; I will be upset when forgetting what I have just heard."*
GLA	LAC	*"Sometimes I get worried before the English listening class because I was afraid to answer questions in English listening class."* *"I worry about the misunderstanding of what the teacher is saying in English listening class, or the unfamiliar topic of the English listening lesson."* *"The most annoying thing in English listening class is to do listening practice in class, and it makes me uneasy."*
	LADU	*"I worry about not understanding the speaker who speaks English fast in a social gathering. Sometime it is hard for me to keep up the speed. If I cannot receive information fluently, I will get worried and I can hear my heartbeat. If I cannot follow the conversation, I will be disappointed with my listening comprehension."* *"Misunderstanding the speaker who speaks English on the phone makes me embarrassed and upset, so I always feel nervous about answering my foreign teachers' phone calls."*
	LAML	*"I like watching English movies and videos on media, and I think it is good for my English learning because I can learn the cultural, know English with different accents, and practice listening. But when I fail to understand what the characters say in movies, I will become nervous and upset. And it makes me frustrated when I have to over-rely on Chinese subtitles to figure out the plot of the movie."* *"I will listen to English news before I go to bed every night. But sometimes I feel uneasy about failing to keep up with the speed of English news."*

Abbreviations: LTA, listening test anxiety; GLA, general listening anxiety; LAC, listening anxiety in classroom; LADU, listening anxiety in daily usage; LAML, listening anxiety in media learning.

college English teachers and foreign teachers, and attending social gatherings. The results indicate that situations that can elicit FLLA extend from traditional classroom and test situations to a great variety of situations in which EFL learners have daily contact with English. Participant 9 felt nervous when receiving information in English. He said:

> I worry about not understanding the speaker who speaks English fast in a social gathering. Sometimes it is hard for me to keep up the speed. If I cannot receive information fluently, I will get worried and I can hear my heartbeat. If I cannot follow the conversation, I will be disappointed with my listening comprehension.

Listening anxiety in media learning referred to a frustrating feeling that EFL learners experienced when they engaged in informal learning on media, such as watching English movies and videos, and listening to English songs and news. The interviews showed that engagement in out-of-school English contact was unintentional language learning or listening practice; however, EFL learners were aware of potential benefits of listening gains from these activities. This finding was in line with the result found by Muñoz and Cadierno (2021) that out-of-school English exposure predicated listening improvement.[6] However, when EFL learners perceived the gap between ideal improvement in listening comprehension and the actual experience of listening comprehension when engaging in out-of-school English contact, they became sensitive to FLLA. Participant 20 said about FLLA she experienced when she watched English movies:

> I like watching English movies and videos on media, and I think it is good for my English learning because I can learn the cultural, know English with different accents, and practice listening. But when I fail to understand what the characters say in movies, I will become nervous and upset. And it makes me frustrated when I have to over-rely on Chinese subtitles to figure out the plot of the movie.

The results of interviews also revealed the relationship between general listening anxiety and listening test anxiety. To a great extent, FLLA elicited in situations like classroom, daily situations, and media learning can be considered as the representatives of general listening anxiety. In this study, general listening anxiety referred to the apprehension that occurs when EFL learners are exposure to a set of anxiety-arousal situations demanding listening skill. On the other

hand, listening test anxiety referred to worry and tension over the comprehension of listening material experienced at a particular moment in test situations. In addition, the findings of the interviews suggest that general listening anxiety affects listening test anxiety. Many of the participants believed listening anxiety they experienced in general situations influenced listening anxiety in tests. For example, Participant 5 described the influence of general listening anxiety on listening test anxiety as follows:

> After a period of listening practice on online APP learning, if I don't gain improvement in my listening, I will put pressure on myself and I feel anxious. The more anxious I feel in listening practice by myself, the more anxious I feel during the listening test.

Although many previous studies dealt with general listening anxiety[11,30] and listening test anxiety,[3,17] respectively, there is no study examining the relationship between them. Our interviews suggest that general listening anxiety has an effect on listening test anxiety, which needs to be further tested.

The Results of EFA

The EFA results showed that Kaiser-Meyer-Olkin (KMO) index for Measure of Sampling Adequacy was 0.913, and Bartlett's test of Sphericity was significant at the 0.001 level. These results indicated the data were appropriate for factor analysis. The 18 items revealed highly related four factors, which supported the interview results. Only one item was discarded for not loading on any factor. The remaining 17 items of the scale are detailed in Table 2. All of the loadings

Table 2 The Scale Items and the Results of Exploratory Factor Analysis of the Scale

Items	Loading			
	Factor 1	Factor 2	Factor 3	Factor 4
LTA1: In the English listening test, once I encounter an unfamiliar key word, I will be so nervous that I cannot keep up with the follow-up content.	0.880			
LTA2: In the English listening test, I am too nervous to remember what I have just heard.	0.973			
LTA3: During the English listening test, I worry about listening to an English passage when I am unfamiliar with the topic.	0.861			
LTA4: In the English listening test, I am afraid that I do not have enough time to think about what I hear.	0.782			
LTA5: Before the English listening test, I always get worried that I cannot understand words or phrases, and this worry affects my listening test performance.	0.694			
LADU1: In daily usage, if people speak English, I get worried when I do not understand what they mean.		0.950		
LADU2: In daily usage, if someone speaks English very fast, I feel very nervous.		0.813		
LADU3: When answering the phone, I get worried when I do not understand what the English speaker is saying.		0.911		
LADU4: In social gatherings, if people around me talk in English, I am very worried about not understanding them.		0.937		
LAML1: When watching videos on social media, if I find that I cannot understand English in the videos, I get worried about my English listening comprehension.				0.587
LAML2: When watching English movies, I worry about not understanding the English dialogue.				0.977
LAML3: When watching English movies, if I find that the content I have heard is inconsistent with the subtitles, I am worried about my English listening comprehension.				0.946
LAML4: When listening to English news, I get more nervous when the news is too fast to keep up.				0.728
LAC1: I am always worried before my English listening class.			0.727	
LAC2: In English listening class, I am nervous when I encounter unfamiliar words.			0.844	
LAC3: In English listening class, I am worried that the topic is unfamiliar to me.			0.683	
LAC4: In English listening class, I feel nervous when I am doing the English listening practice.			0.951	

Abbreviations: LTA, listening test anxiety; LADU, listening anxiety in daily usage; LAML, listening anxiety in media learning; LAC, listening anxiety in classroom.

met the cutoff values proposed by Hair et al[50] (See Table 2). Finally, four distinct factors with 17 items emerged from the EFA, accounting for 76.827% of the variance. EFA generated four factors, named *Listening Test Anxiety* (LTA), *Listening Anxiety in Daily Usage* (LADU), *Listening Anxiety in Classroom* (LAC), and *Listening Anxiety in Media Learning* (LAML).

The Results of CFA

A CFA was performed on the second batch of the data to assess the four-component structure yielded via the EFA. It is suggested that the root mean square error of approximation (RMSEA) should be in the range between 0.05 and 0.08 (with the lower bound of 90% confidence interval less than 0.05 and the upper bound less than 0.10), and standardized root mean square residual (SRMR) should be less than 0.08.[51] The results of the fit indices χ^2 = 201.503, df = 113, p < 0.001, χ^2/df = 1.783, RMSEA = 0.062 [0.048, 0.075], SRMR = 0.041, together with the value of comparative fit index (CFI) 0.966, suggested that the model was reasonable fit, meeting general guidelines by Hair et al.[50]

Convergent validity and discriminant validity tests were conducted to assess whether the model had an adequate level of validity. Convergent validity tests were examined by assessing standardized factor loadings, average variance extracted (AVE) values, and construct reliability (CR). According to the recommended thresholds suggested by Hair et al, [50] the results showed that standardized factor loadings were higher than 0.7 and significant (p < 0.001); the AVEs that ranged from 0.598 to 0.781 were all greater than the recommended 0.5 threshold, and CR values (ranging from 0.882 to 0.934) were larger than 0.7 threshold (See Table 3). In CFA, discriminant validity is to estimate the extent to which a construct is different from another, which can be tested by Fornell-Larcker Criterion.[50] Fornell-Larcker Criterion is the positive square root of the AVE for the each construct that is higher than the highest correlation with constructs. Table 4 details the results of discriminant validity, indicating that the four dimensions were distinct from each other. In sum, the results of convergent validity and discriminant validity tests suggest the scale was of good construct validity.

Theoretically, LADU, LAML, LAC, and general listening anxiety (GLA) constituted a second-order construct that was a reflective-reflective type of construct. Based on the method of identifying and testing of a higher-order construct

Table 3 Results of Confirmatory Factor Analysis

Constructs	Items	UFL	*SE*	*z*	SFL	CR	Cronbach's α	AVE
LTA	LTA1	1.011***	0.084	12.020	0.745	0.882	0.881	0.598
	LTA2	1.055***	0.082	12.918	0.784			
	LTA3	1.129***	0.086	13.109	0.791			
	LTA4	0.939***	0.076	12.429	0.762			
	LTA5	1.047***	0.081	12.905	0.784			
LADU	LADU1	1.186***	0.078	15.166	0.856	0.934	0.934	0.781
	LADU2	1.138***	0.074	15.380	0.863			
	LADU3	1.117***	0.067	16.724	0.907			
	LADU4	1.198***	0.072	16.724	0.907			
LAML	LAML1	0.953***	0.076	12.573	0.765	0.885	0.883	0.658
	LAML2	1.013***	0.078	12.945	0.780			
	LAML3	1.104***	0.075	14.771	0.851			
	LAML4	1.053***	0.072	14.637	0.845			
LAC	LAC1	1.204***	0.084	14.310	0.832	0.895	0.894	0.682
	LAC2	1.202***	0.077	15.614	0.880			
	LAC3	1.156***	0.079	14.728	0.847			
	LAC4	1.050***	0.088	11.982	0.737			

Note: *** p < 0.001.
Abbreviations: LTA, listening test anxiety; LADU, listening anxiety in daily usage; LAML, listening anxiety in media learning; LAC, listening anxiety in classroom; UFL, unstandardized factor loadings; SE, standard error; SFL, standardized factor loadings; CR, composite reliability; AVE, average variance extracted.

Table 4 Results of Fornell-Larcker Criterion Test for Confirmatory Factor Analysis

	LAC	**LAML**	**LADU**	**LTA**
LAC	**0.826**			
LAML	0.726	**0.811**		
LADU	0.597	0.746	**0.884**	
LTA	0.689	0.628	0.584	**0.773**

Note: The bolded values are the positive square roots of the AVE for the corresponding constructs.
Abbreviations: LAC, listening anxiety in classroom; LAML, listening anxiety in media learning; LADU, listening anxiety in daily usage; LTA, listening test anxiety.

proposed by Doll et al, [52] four plausible alternative models composed of LADU, LAML, LAC, and GLA were compared and tested the fit of each hypothesized model. The results suggest that both model 3 and model 4 show better model-data fit than model 1 and model 2. The results of alternative models for the FLLA scale are presented in Table 5. Based on guidelines by Doll et al, [52] model 4 with three first-order factors (LADU, LAML, and LAC) and one second-order (GLA) provided higher validity than other alternative models, which is of greater theoretical interest than other models.

Study 2
Research Hypotheses

In study 2, we examined the relationship between general listening anxiety, listening test anxiety, and self-perceived listening performance through structural equation modeling (SEM). Previous research has yielded mixed results on the relationship between FLLA and listening performance. For example, several studies have found a significant negative relationship between FLLA and listening achievement;[28,53] whilst some studies have not detected any significant correlation between FLLA and listening performance.[10,54] These incongruent outcomes may partially attribute to unclear distinctions between different types of FLLA. Given CDST has endeavored to explain the stability and instability of behaviors, which provides strong evidence for the distinction between general listening anxiety and listening test anxiety. According to this theory, when the system is more unstable, situational factors can have a great influence.[55] It suggests that listening test anxiety tends to be more heavily influenced by external situational factors (such as the high-stakes of test situations) than general FLLA does. Thus, we hypothesized that general listening anxiety and listening test anxiety have different effects on self-perceived listening performance. Specifically, the following hypotheses were formed:

H1: General listening anxiety has a positive effect on listening test anxiety.

H2: General listening anxiety has a negative effect on self-perceived listening performance.

H3: Listening test anxiety has a negative effect on self-perceived listening performance.

H4: General listening anxiety, mediated through listening test anxiety, influences self-perceived listening performance.

Table 5 Alternative Models for the Foreign Language Listening Anxiety Scale

	χ^2 (df)	χ^2 / df	RMSEA [90% CI]	CFI	SRMR
Model 0. Null model	1976.801 (66)	29.952	0.374 [0.360, 0.388]	0.000	0.513
Model 1. 1 First-order factor	541.208 (54)	10.022	0.209 [0.193, 0.225]	0.745	0.106
Model 2. 3 First-order factors (Uncorrelated)	337.765 (54)	6.255	0.159 [0.143, 0.176]	0.852	0.383
Model 3. 3 First-order factors (Correlated)	90.866 (51)	1.782	0.062 [0.040, 0.082]	0.979	0.038
Model 4. 3 First-order factors Second-order factor	90.866 (51)	1.782	0.062 [0.040, 0.082]	0.979	0.038

Abbreviations: RMSEA, root mean square error of approximation; CFI, comparative fit index; SRMR, standardized root mean square residual.

Based on previous research, it was found that the intensity of FLLA varied considerably due to different majors.[39] For the situation-specific perspective on FLLA discussed in this study, we were more concerned about the differences in general listening anxiety between English and non-English major students; thereby we put forward the following hypothesis:

H5: There is a significant difference in general listening anxiety between English and non-English major students.

Besides the possible differences in general listening anxiety, there may be differences in the strength of the impacts of general listening anxiety and listening test anxiety on listening performance between different majors, even though there is few research involving this issue. Therefore, we put forward the following research hypotheses:

H6: There is a significant difference in the impact of general listening anxiety on listening test anxiety between English and non-English major students.

H7: There is a significant difference in the effect of general listening anxiety on self-perceived listening performance between English and non-English major students.

H8: There is a significant difference in the effect of listening test anxiety on self-perceived listening performance between English and non-English major students.

Participants and Procedures

As mentioned in the literature review, to collect the feelings of self-reported listening anxiety, the scale should be distributed immediately after the listening test. In this study, the survey was conducted right after the listening course tests. We recruited all the senior English majors in a university in a city in central China as the participants (not the same cohort of participants in Study 1). A total of 163 English majors took the exam, and 152 participants voluntarily took part in the survey and finished the questionnaire. Among the included 152 participants, 13 (9%) were males and 139 (91%) were females. The participants ranged from 20 to 24 years of age ($M = 21.28$, $SD = 0.743$). All the participants had learned English as a foreign language from 10 to 14 years, and none of them had experience living in any English-speaking country. A total of 228 non-English majors participated in the survey and completed the questionnaire. Among them, 68 (30%) were males and 160 (70%) were females, and they were aged between 18 and 23 ($M = 19.55$, $SD = 0.793$). All these participants, prior to college, had learned English as a second language for at least six years. At the college, the College English course is compulsory for these non-English major students for the first- and second-year of study.

Instruments

Foreign Language Listening Anxiety Scale

We employed the newly developed FLLA scale to measure listening anxiety. The scale consisted of 17 items that comprised four factors evaluated by EFA and CFA above. The items were answered on a 7-point Likert scale, ranging from 7 (strongly agree) to 1 (strongly disagree).

Self-Perceived Performance Assessment

The self-perceived performance referred to test takers' self-assessments of the current listening test performance, and the full score of the self-assessment was 100. This self-assessment score was not the estimation of the actual score in the test, but the self-assessment of the performance of this test that whether they have outperformed or underperformed in this test.

Data Analysis

SmartPLS 3.3.2[56] was used for data analysis, because it has more advantages in dealing with a single item structure and small samples than covariance-based structural equation modeling, and it can better deal with high-order models. The data processing involved the following three procedures. First, measurement model testing was used to evaluate whether

the reliability and validity of construct measurements met the requirements. The two-stage method[50] was employed for the analysis, because general listening anxiety in the model was a second-order construct that consisted of three first-order constructs as a new reflective-reflective construct. In the first stage, the repeated indicator approach was used to examine the measurement models of the first-order constructs (LAC, LADU, and LAML). In the second stage, the latent variable scores of the three first-order constructs were obtained (LAC_LV, LADU_LV, LAML_LV), which served as the indicators of the higher second-order construct (GLA).[50] Structural model testing and mediating effect testing were then conducted to test the hypotheses of this study.

Results
Assessment of Measurement Model

The assessment of the measurement model in PLS-SEM involves testing internal consistency reliability, convergent validity, and discriminant validity by judging the indicators of factor loadings, composite reliability (CR), Cronbach's α, average variance extracted (AVE), and Fornell-Larcker Criterion. The results showed that the measurement models had good internal consistency, reliability, and convergence, meeting guidelines by Hair et al.[50] The specific statistics of factor loading, CR, Cronbach's α, and AVE are provided in Table 6. The Fornell-Larcker criterion test indicates that the measurement model has good discriminant validity (see Table 7 and Table 8).

Measurement Model Invariance Testing

The measurement variance of composite models (MICOM) was used to assess the measurement model invariance. Based on the guideline of Hair et al, [50] a three-step analysis was performed and the results showed that all the three constructs

Table 6 Results of the Test of Measurement Models

Constructs		Items	Loading		CR		Cronbach's α		AVE	
			>0.7		>0.7		>0.7		>0.5	
			EMS	Non-EMS	EMS	Non-EMS	EMS	Non-EMS	EMS	Non-EMS
First-order constructs	LAC	LAC1	0.860	0.875	0.925	0.938	0.891	0.912	0.755	0.791
		LAC2	0.866	0.900						
		LAC3	0.907	0.910						
		LAC4	0.841	0.872						
	LADU	LADU1	0.913	0.898	0.951	0.948	0.932	0.926	0.830	0.819
		LADU2	0.910	0.910						
		LADU3	0.910	0.911						
		LADU4	0.911	0.902						
	LAML	LAML1	0.843	0.839	0.911	0.926	0.870	0.894	0.720	0.758
		LAML2	0.866	0.881						
		LAML3	0.879	0.880						
		LAML4	0.804	0.883						
	LTA	LTA1	0.751	0.809	0.902	0.918	0.864	0.888	0.648	0.692
		LTA2	0.830	0.830						
		LTA3	0.774	0.845						
		LTA4	0.811	0.820						
		LTA5	0.855	0.853						
Second-order construct	GLA	LAC_LV	0.856	0.863	0.898	0.924	0.830	0.877	0.746	0.803
		LADC_LV	0.851	0.894						
		LAML_LV	0.883	0.929						

Abbreviations: LAC, listening anxiety in classroom; LADU, listening anxiety in daily usage; LAML, listening anxiety in media learning; LTA, listening test anxiety; GLA, general listening anxiety; EMS, English major students; Non-EMS, non-English major students; LV, latent variables scores; CR, composite reliability; AVE, average variance extracted.

Table 7 Results of Fornell-Larcker Criterion Test of the First-Order Constructs

	EMS			Non-EMS		
	LAC	**LADU**	**LAML**	**LAC**	**LADU**	**LAML**
LAC	**0.869**			**0.889**		
LADU	0.545	**0.911**		0.608	**0.905**	
LAML	0.647	0.664	**0.849**	0.748	0.761	**0.871**

Notes: The bolded values are the positive square roots of the AVE for the corresponding constructs.
Abbreviations: LAC, listening anxiety in classroom; LADU, listening anxiety in daily usage; LAML, listening anxiety in media learning; EMS, English major students; Non-EMS, non-English major students.

Table 8 Results of Fornell-Larcker Criterion Test of the Second-Order Constructs

	EMS		Non-EMS	
	GLA	**LTA**	**GLA**	**LTA**
GLA	**0.864**		**0.896**	
LTA	0.621	**0.805**	0.728	**0.832**

Notes: The bolded values are the positive square roots of the AVE for the corresponding constructs.
Abbreviations: GLA, general listening anxiety; LTA, listening test anxiety; EMS, English major students; Non-EMS, non-English major students.

established configural and compositional invariance but the constructs did not have equal mean values and variances, indicating that partial measurement invariance was confirmed. The partial measurement invariance suggests that the comparison of differences between groups in models can be analyzed but data of different groups cannot be pooled together. The results of the equal mean assessment also showed that there were significant differences in general listening anxiety intensity between the two groups, suggesting that H5 was supported. Specifically, the findings revealed that non-English major students' general listening anxiety was higher than that of English major students. Results of invariance measurement testing are presented in Table 9.

Structural Model Testing

To test the relationships among the constructs, bootstrapping with 5000 subsamples was run to assess the significance of path coefficients. The results showed that general listening anxiety had a significant effect on listening test anxiety, and listening test anxiety had a significant effect on self-perceived listening test performance. However, no significant relationship between general listening anxiety and self-perceived listening test performance was found. The summary of the indices for the testing of the structure model in Table 10 suggests that H1 and H3 were supported but H2 was rejected. In terms of evaluating mediation analysis, results showed that the mediation role of listening test anxiety was established (see Table 10 for details). According to the mediation analysis method of Zhao et al, [57] it was found that listening test anxiety played a full mediation effect on the relationship between general listening anxiety and self-perceived listening performance, indicating that H4 was supported. In terms of coefficient of determination, Chin suggested that R^2 values higher than 0.67, 0.33, and 0.19 indicate a strong, moderate, and weak coefficient of determination, respectively.[58] As summarized in Figures 1 and 2, the models of the two groups moderately explained the variation of listening test anxiety ($R^2_{EMS} = 0.385$, $R^2_{non-EMS} = 0.530$) and weakly explained self-perceived listening test performance ($R^2_{EMS} = 0.208$, $R^2_{non-EMS} = 0.225$). In terms of model fitting, the value of SRMR of the two groups was 0.072 and 0.071, respectively, which is less than the recommended cut-off value of 0.08,[50] indicating that the model fitting was acceptable.

Table 9 Results of Invariance Measurement Testing

Constructs	Compositional Invariance (Correlation = 1)		Partial Measurement Invariance	Equal Mean Assessment			Equal Variance Assessment			Full Measurement Invariance
	C = 1	Confidence Interval		Differences (EMS-Non-EMS)	Confidence Interval	Equal	Difference (EMS-Non-EMS)	Confidence Interval	Equal	
GLA	0.999	[0.999, 1.000]	Yes	−0.298	[−0.201, 0.207]	No	0.014	[−0.311, 0.299]	Yes	No
LTA	0.999	[0.998, 1.000]	Yes	−0.143	[−0.207, 0.203]	Yes	−0.099	[−0.319, 0.297]	Yes	Yes
SPLP	1.000	[1.000, 1.000]	Yes	1.056	[−0.207, 0.209]	No	−0.689	[−0.329, 0.306]	No	No

Abbreviations: GLA, general listening anxiety; LTA, listening test anxiety; SPLP, self-perceived listening performance.

Table 10 Results of the Test of Structural Models

Hypothesis	Relationships/ Indirect Effects	Path Coefficient		SD		t		R^2		Supported
		EMS	Non-EMS	EMS	Non-EMS	EMS	Non-EMS	EMS	Non-EMS	
H1	GLA -> LTA	0.621***	0.728***	0.057	0.037	10.898	19.627	0.385	0.530	Yes
H2	GLA -> SPLP	−0.170ns	−0.057ns	0.107	0.082	1.586	0.689			No
H3	LTA -> SPLP	−0.331***	−0.431***	0.093	0.082	3.572	5.235	0.208	0.225	Yes
H4	GLA -> LTA -> SPLP	−0.205**	−0.314***	0.060	0.065	3.431	4.864			Yes

Note: **$p < 0.01$, ***$p < 0.001$
Abbreviations: ns, not significant; GLA, general listening anxiety; LTA, listening test anxiety; SPLP, self-perceived listening performance; EMS, English major students; Non-EMS, non-English major students; SD, standard deviation.

Multi-Group Analysis

As detailed in Table 11, the results of the multi-group analysis showed that there was no significant difference between English and non-English major students in all three paths of the model. These findings showed that the structures were the same in the two models. Thus, H6, H7, and H8 were rejected.

Discussion

The present study has taken a closer look at the distinction between general listening anxiety and listening test anxiety and the relationship between these listening anxieties and listening performance. The results of this research revealed that general anxiety-arousal situations involved classroom, English daily usage contexts, and media learning settings. The componential factors of foreign language listening anxiety included two factors, one having to do with general listening anxiety; the other with listening test anxiety. Specifically, general listening anxiety was typically represented by three sub-components: listening anxiety in classroom, listening anxiety in daily usage, and listening anxiety in media learning. The results also found that general listening anxiety had no direct effect on self-perceived listening performance; while listening test anxiety negatively affected self-perceived listening performance, and meanwhile listening test anxiety played a full mediation role in the relationship between general listening anxiety and self-perceived listening anxiety. In addition, a significant difference in the level of general listening anxiety was found between English and non-English
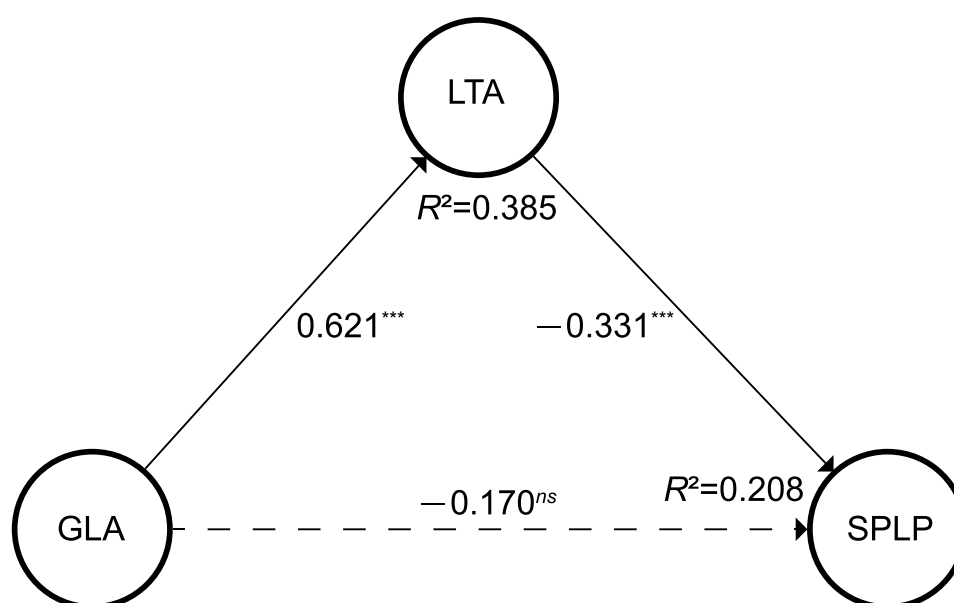

Figure 1 Results of the structural models analysis for English major group.
Note: ***$p < 0.001$.
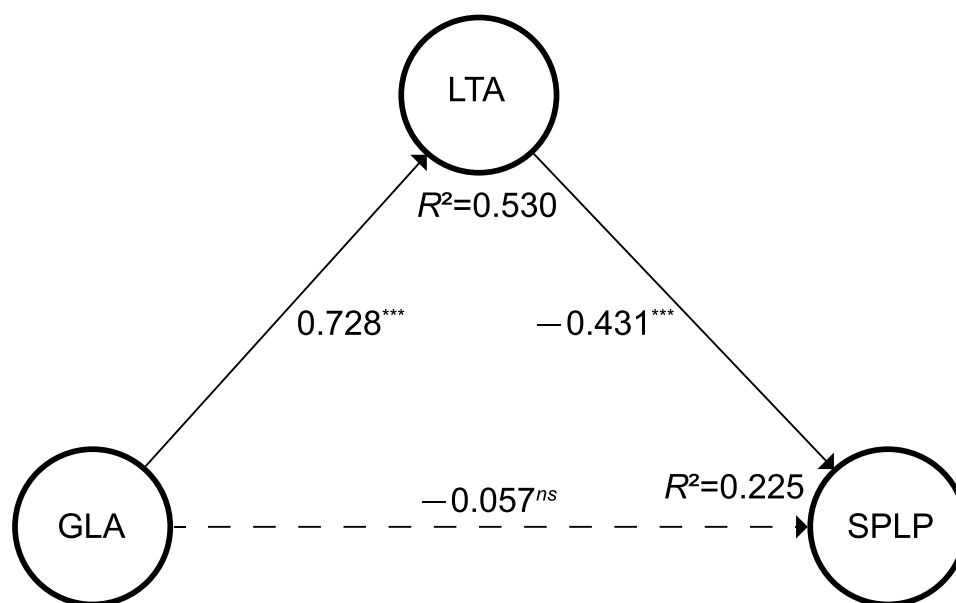**Abbreviation**: ns, not significant.

Figure 2 Results of the structural models analysis for non-English major group.
Note: ***p < 0.001
Abbreviation: ns, not significant.

majors; however, no significant difference in anxiety-performance linkage was found between the two groups from the multi-group analysis.

With regard to research question one concerning the measurement of FLLA, the results of the qualitative study found the emergence of four factors of FLLA: listening test anxiety, listening anxiety in daily usage, listening anxiety in classroom, and listening anxiety in media learning. And further the quantitative study has provided robust evidence for the distinctness of two types of listening anxiety: general listening anxiety and listening test anxiety. Findings across the studies supported the separation of listening test anxiety from general listening anxiety. These findings aligned with the results of studies that test anxiety should be independent of language anxiety,[15] yet diverged from Horwitz et al claim that test anxiety was one of the componental factors of general foreign language anxiety.[12] Differentiating between listening test anxiety and general listening anxiety has expanded knowledge of the phenomenology of listening anxiety and helped identify potentially stressful situational characteristics as sources of listening anxiety. Furthermore, the empirical results of this research have supported the theoretical distinction between general listening anxiety and listening test anxiety under the framework of CDST. The CDST postulates that attractor state describes a dynamic system moving from one state to another as a result of the interactions between the individual factors and the nature of the situation.[18] The nature of a listening test situation that characterizes as the time pressure for receiving, decoding, and processing transient spoken information[59] and highly emphasizes task-oriented performance[28] may act as a disturbing force that can change the original stable attractor state into an increased intensity of tension and worry.

Table 11 Results of Multi-Group Analysis

Hypothesis	Relationships	Path Coefficient Difference (EMS - Non-EMS)	Supported
H6	GLA -> LTA	-0.107^{ns}	No
H7	GLA -> SPLP	-0.114^{ns}	No
H8	LTA -> SPLP	0.100^{ns}	No

Abbreviations: ns, not significant; GLA, general listening anxiety; LTA, listening test anxiety; SPLP, self-perceived listening performance.

In addition, evidence combined with the semi-structured interviews and qualitative analysis revealed that general anxiety-arousal situations included classroom, daily communication contexts, and media learning settings. This finding was not consistent with the result from previous research that general language anxiety was confined to classroom.[1,17] It indicates that new learning contexts can arouse unique general listening anxiety, which was congruent with the finding that general language anxiety was affected by contextual influences.[16,18,19] The result of the existence of one second-order factor and three first-order factors suggests that there exists a relatively stable general listening anxiety across several situations. This result concurred with the study by Hwang et al in that general English anxiety cannot be reduced by merely providing more English exposure in mobile-assisted language learning contexts.[60] The results also suggested that general listening anxiety used to situated in the static classroom[17] has evolved into various context-dependent situations. With the development of technology and media, an array of out-of-school learning activities expanded language learning from the classroom context to daily and online learning settings;[61] thereby it was necessary to explore language leaners' emotional experiences beyond the classroom.[24]

In respondence to research question two about the anxiety-and-performance relationship, results from the structural equation modeling confirmed different effects of listening anxieties on listening performance. Results found that only listening test anxiety negatively affected self-perceived listening performance ($R^2_{EMS} = 0.208$). This finding was consistent with the quantitative research that test anxiety negatively influenced performance.[62] However, general listening anxiety had no direct effect on self-perceived listening performance. Combined with the results from qualitative analysis, the empirical results of this study offered greater support for the need to separate listening test anxiety from general listening anxiety.[8] This finding may help explain why no significant relationship between listening anxiety and listening test performance was detected in previous research,[10] and why intervention in extensive listening failed to lower listening anxiety.[63] Because unawareness of the distinctions between general listening anxiety and listening test anxiety causes the mixed use of these two listening anxieties to tap into the causal relationship between general listening anxiety and listening test performance; however, general listening anxiety does not have a direct effect on listening test performance, which was not congruent with the findings of by Elkhafaifi (2005),[30] in that general listening anxiety was a strong predictor of listening test performance. The findings of this study indicate that a critically important distinction between different types of listening anxiety, which was overlooked in previous research, is a particularly useful path for anxiety-and-performance research.

In terms of the structural framework of listening anxieties and listening performance, the SEM results provided empirical evidence for the full mediation role of listening test anxiety on the relationship between general listening anxiety and self-perceived listening performance. Notably, this finding was inconsistent with previous research[17,27,53] in that empirical evidence was not found for the direct adverse impact of general listening anxiety on listening performance. It suggests that the separation of general listening anxiety and listening test anxiety may provide precise and accurate evidence for the detection of the relation between listening anxiety and performance, because it has been proved that general listening anxiety indirectly affected listening test performance through the full mediation effect of listening test anxiety. The result also suggests that the inconsistency in argumentation in previous research[3,9,10] concerning the relationship between FLLA and performance can be resolved by separating different types of FLLA and then exploring different effects of listening anxieties on performance. In addition, this empirical finding may support the theoretical assumptions in CDST that attractor states are assumed as a result of self-organizing, and external sources from environment can disturb the stability of the system.[18] General listening anxiety can self-organize and self-regulate to settle in a relatively stable attractor state, which will not directly influence listening test performance. However, the test situation can perturb the stable attractor state into an anxious state that in turn directly influences listening test performance. Nevertheless, caution is required when generalizing from the main findings of this study, because learners with different cultural backgrounds may have multiple types of out-of-school target language exposures, which might contribute differently to the effect of general listening anxiety on listening test anxiety.

With regard to research question three concerning the difference in general listening anxiety between English and non-English major students, the results of MICOM of this study showed that a significant difference between the two groups was found, with non-English majors' general listening anxiety higher than that of English majors. To explain this finding, two considerations are put forward. One possibility is that language proficiency plays an important role in the effect of general listening anxiety on learners with different majors. Given the fact that the language proficiency of

English majors is higher than that of non-English majors, and the settled issue that proficiency is also regarded as a cause of anxiety,[19] it is possible that English majors are not as strongly influenced by general listening anxiety as their non-English major peers. A second possibility is that the amount of out-of-school exposure to English listening may have effects on the general listening anxiety between different majors. Non-English majors may be less frequently exposed to out-of-school incidental language learning settings. It can be supported by the results gained from the semi-structure interviews that non-English majors were not as sensitive as English majors in generating anxiety-arousal situations; in other words, when non-English majors were prompted to identify possible anxiety-arousal situations, they responded that they experienced listening anxiety in these situations to a certain degree. Insensitively realizing the existence of these situations may suggest that non-English majors are not frequently exposed to out-of-school English contact situations. However, results obtained from the multi-group analysis showed that English majors did not differ from non-English majors in the effect of general listening anxiety on listening test anxiety, in the effect of general listening anxiety on self-perceived listening test performance, and in the effect of listening test anxiety on self-perceived listening test performance. The results revealed that the structure of listening anxiety affecting listening performance proposed in this study was relatively stable between English and non-English majors. For the two groups, the mechanism of anxiety-and-performance relationship was consistent.

Conclusions and Implications

The present research investigated foreign language listening anxiety both in and beyond the classroom setting and, for the first time, verified the component structure of foreign language listening anxiety. The analysis of the newly developed scale provided evidence in support of a two-component structure of foreign language listening anxiety: general listening anxiety and listening test anxiety. Further structural equation modeling analyses showed that general listening anxiety positively influenced listening test anxiety; listening test anxiety negatively influenced self-perceived listening test performance, and general listening anxiety indirectly affected self-perceived listening performance through the full mediation effect of listening test anxiety. For the group comparison analysis, results showed that non-English majors' general listening anxiety was higher than that of English majors; however, the two groups did not show a difference in the intensity difference of the impact of general listening anxiety on listening test anxiety, and in the intensity differences of the impacts of listening anxieties on self-perceived listening performance.

Several implications for teachers can be drawn from this study. First, teachers should recognize the increasing importance of foreign language listening practice in out-of-school settings because lowering listening anxiety in general situations may effectively reduce listening anxiety in tests and improve test performance accordingly. Second, there is a need to reduce listening test anxiety as it influences listening test performance directly. Thereby, teachers should help students to regulate their worry and apprehension with effective strategies.

This research expands upon the findings of previous listening anxiety studies; however, it is not without limitations. One is the unbalanced proportion of male and female samples, because the overwhelming majority of English majors in universities in mainland China are female. In future studies, the samples can be enlarged with more male participants participating in the research. Second, in this research we only focused on English as a foreign language; however, we do not have evidence to show whether the newly developed foreign language listening anxiety scale is applicable to other types of foreign languages. Future research needs to validate the structure of foreign language listening anxiety in learners with various learning contexts. Third, although the results of the present research have shown that general listening anxiety had a significant effect on listening test anxiety, other factors may have a direct influence on listening test anxiety. For example, Chang (2008) has found that worrying about test difficulty was the source of listening test anxiety.[17] Therefore, future studies can deepen the research by explicitly identifying various factors that influence listening test anxiety.

Ethical Approval

This study was conducted based on the guidelines of the Declaration of Helsinki, and the guidelines of Measures for the Ethical Review of Biomedical Research Involving Humans released by the National Health and Family Planning Commission (China), and it was approved by the Academic Committee of the School of Journalism and Communication, Wuhan University.

Informed Consent Statement

Informed consent was obtained from all participants involved in the study.

Acknowledgments

We would like to express sincerest gratitude to all the participants who supported this study.

Funding

Disclosure

Dr Suhe Ji reports grants from Central China Normal University, during the conduct of the study. The authors report no other conflicts of interest in this work.

References

1. Zhang X. Foreign language anxiety and foreign language performance: a meta-analysis. *The Mod Lang J*. 2019;103(4):763–781. doi:10.1111/modl.12590
2. Jin Y, Dewaele JM, MacIntyre PD. Reducing anxiety in the foreign language classroom: a positive psychology approach. *System*. 2021;101:102604. doi:10.1016/j.system.2021.102604
3. In'nami Y. The effects of test anxiety on listening test performance. *System*. 2006;34(3):317–340. doi:10.1016/j.system.2006.04.005
4. Xie Q. Does test preparation work? Implications for score validity. *Lang Assess Q*. 2013;10(2):196–218. doi:10.1080/15434303.2012.721423
5. Jiang Y, Dewaele JM. The predictive power of sociobiographical and language variables on foreign language anxiety of Chinese university students. *System*. 2020;89. doi:10.1016/j.system.2020.102207
6. Muñoz C, Cadierno T. How do differences in exposure affect English language learning? A comparison of teenagers in two learning environments. *Stud Second Lang Learn Teach*. 2021;11(2):185–212. doi:10.14746/ssllt.2021.11.2.2
7. Lindgren E, Muñoz C. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *Int J Multiling*. 2013;10(1):105–129. doi:10.1080/14790718.2012.679275
8. Ji S, Qin X, Li K. A systematic review of foreign language listening anxiety: focus on the theoretical definitions and measurements. *Front Psychol*. 2022;13(June):1–17. doi:10.3389/fpsyg.2022.859021
9. Kim J. Foreign language listening anxiety: a study of Korean students learning English; 2000.
10. Liu M. Interrelations between foreign language listening anxiety and strategy use and their predicting effects on test performance of high- and low-proficient Chinese university EFL learners. *Asia Pacific Educ Res*. 2016;25(4):647–655.
11. Zhang X. Foreign language listening anxiety and listening performance: conceptualizations and causal relationships. *System*. 2013;41(1):164–177. doi:10.1016/j.system.2013.01.004
12. Horwitz EK, Horwitz MB, Cope J. Foreign language classroom anxiety. *Mod Lang J*. 1986;70(2):125–132. doi:10.5840/ajs2012281-29
13. Horwitz EK. Foreign and second language anxiety. *Lang Teach*. 2010;43(2):154–167. doi:10.1017/S026144480999036X
14. MacIntyre PD, Gardner RC. Methods and results in the study of anxiety and language learning: a review of the literature. *Lang Learn*. 1991;41(1):85–117. doi:10.1111/j.1467-1770.1991.tb00677.x
15. Aida Y. Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: the case of students of Japanese. *Mod Lang J*. 1994;78(2):155–168. doi:10.1111/j.1540-4781.1994.tb02026.x
16. Dewaele JM, Alfawzan M. Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Stud Second Lang Learn Teach*. 2018;8(1):21–45. doi:10.14746/ssllt.2018.8.1.2
17. Chang ACS. Sources of listening anxiety in learning English as a foreign language. *Percept Mot Skills*. 2008;106:21–34. doi:10.1108/eb003506
18. Dörnyei Z, MacIntyre PD, Henry A. *Motivational Dynamics in Language Learning*. Multilingual Matters; 2015.
19. MacIntyre PD. An overview of language anxiety research and trends in its development. In: Gkonou C, Daubney M, Dewaele JM, editors. *In New Insights into Language Anxiety: Theory, Research and Educational Implications*. Bristol: Multilingual Matters; 2017:11–30.
20. Newmark CS, Wheeler D, Newmark L, Stabler B. Test-induced anxiety with children. *J Pers Assess*. 1975;39(4):409–413. doi:10.1207/s15327752jpa3904_15
21. Liu M, Xu H. Testing effects of foreign language listening anxiety on Chinese university students' *English* listening test performance. *Front Psychol*. 2021;12(July):1–7. doi:10.3389/fpsyg.2021.701926
22. Hu X, Zhang X, McGeown S. Foreign language anxiety and achievement: a study of primary school students learning English in China. *Lang Teach Res*. 2021. doi:10.1177/13621688211032332
23. Dewaele JM, Petrides KV, Furnham A. Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: a review and empirical investigation. *Lang Learn*. 2008;58(4):911–960. doi:10.1111/j.1467-9922.2008.00482.x
24. Ross AS, Rivers DJ. Emotional experiences beyond the classroom: interactions with the social world. *Stud Second Lang Learn Teach*. 2018;8(1):103–126. doi:10.14746/ssllt.2018.8.1.5
25. Chang ACS, Millett S. The effect of extensive listening on developing L2 listening fluency: some hard evidence. *ELT J*. 2014;68(1):31–40. doi:10.1093/elt/cct052

26. Pavia N, Webb S, Faez F. Incidental vocabulary learning through listening to songs. *Stud Second Lang Acquis*. 2019;41(4):745–768. doi:10.1017/S0272263119000020

27. Chow BW-Y, Chiu HT, Wong SWL. Anxiety in reading and listening English as a foreign language in Chinese undergraduate students. *Lang Teach Res*. 2018;22(6):719–738.

28. Fathi J, Derakhshan A, Torabi S. The effect of listening strategy instruction on second language listening anxiety and self-efficacy of Iranian EFL learners. *SAGE Open*. 2020;10(2). doi:10.1177/2158244020933878

29. Mills N, Pajares F, Herron C. Self-efficacy of college intermediate French students: relation to achievement and motivation. *Lang Learn*. 2007;57 (3):417–442.

30. Elkhafaifi H. Listening comprehension and anxiety in the Arabic language classroom. *Mod Lang J*. 2005;89(2):206–220. doi:10.1111/j.1540-4781.2005.00275.x

31. Lee SP. Computer-detected attention affects foreign language listening but not reading performance. *Percept Mot Skills*. 2016;123(1):33–45. doi:10.1177/0031512516657337

32. Yang X. Intentional forgetting, anxiety, and EFL listening comprehension among Chinese college students. *Learn Individ Differ*. 2010;20(3):177–187.

33. Teimouri Y, Goetze J, Plonsky L. Second language anxiety and achievement a meta-analysis. *Stud Second Lang Acquis*. 2019;41(2):363–387. doi:10.1017/S0272263118000311

34. Brown AV, Plonsky L, Teimouri Y. The use of course grades as metrics in L2 research: a systematic review. *Foreign Lang Ann*. 2018;763–778. doi:10.1111/flan.12370

35. Liu M, Yuan R. Changes in and effects of foreign language classroom anxiety and listening anxiety on Chinese undergraduate students' English proficiency in the COVID-19 context. *Front Psychol*. 2021;12. doi:10.3389/fpsyg.2021.670824

36. Kalat JW. *Introduction to Psychology*. Wadsworth Publishing; 2010.

37. Cakici D. The correlation among EFL learners' test anxiety, foreign language anxiety and language achievement. *English Lang Teach*. 2016;9 (8):190–203. doi:10.5539/elt.v9n8p190

38. Yamauchi Y. Multifaceted images of Japanese EFL learners' listening anxiety attributable to instructional factors. *Int J Curric Dev Pract*. 2014;16 (1):1–11. doi:10.18993/jcrdaen.16.1_1

39. Kimura H. Foreign language listening anxiety: its dimensionality and group differences. *JALT J*. 2008;30(2):173–196. doi:10.37546/jaltjj30.2-2

40. Ngo H, Spooner-Lane R, Mergler A. A comparison of motivation to learn English between English major and non-English major students in a Vietnamese university. *Innov Lang Learn Teach*. 2017;11(2):188–202. doi:10.1080/17501229.2015.1094076

41. Warden CA, Lin HJ. Existence of integrative motivation in an Asian EFL setting. *Foreign Lang Ann*. 2000;33(5):535–545. doi:10.1111/j.1944-9720.2000.tb01997.x

42. Zhang D. Affective cognition of students' autonomous learning in college English teaching based on deep learning. *Front Psychol*. 2022;12:1–9. doi:10.3389/fpsyg.2021.808434

43. Lee JS, Lee K. Perceptions of English as an international language by Korean English-major and non-English-major students. *J Multiling Multicult Dev*. 2019;40(1):76–89. doi:10.1080/01434632.2018.1480628

44. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm Policy Ment Heal Ment Heal Serv Res*. 2015;42(5):533–544. doi:10.1007/s10488-013-0528-y.Purposeful

45. Francis JJ, Johnston M, Clare R, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychol Health*. 2010;25(10):1229–1245. doi:10.1080/08870440903194015

46. Hesse-Biber S. Gender differences in psychosocial and medical outcomes stemming from testing positive for the BRCA1/2 genetic mutation for breast cancer: an explanatory sequential mixed methods study. *J Mix Methods Res*. 2018;12(3):280–304. doi:10.1177/1558689816655257

47. Yan JX, Horwitz EK. Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: a qualitative analysis of EFL learners in China. *Lang Learn*. 2008;58(1):151–183. doi:10.1111/j.1467-9922.2007.00437.x

48. Corp IBM. IBM SPSS statistics for windows, version 20.0; 2011.

49. Arbuckle JL. Amos (Version 20.0); 2011.

50. Hair JF, Hult GTM, Ringle CM, Sarstedt M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Sage publications; 2014.

51. Kline RB. *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press; 2016.

52. Doll WJ, Xia W, Torkzadeh G. A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Q*. 1994;18(4):453–461.

53. Bang S, Hiver P. Investigating the structural relationships of cognitive and affective domains for L2 listening. *Asian Pacific J Second Foreign Lang Educ*. 2016;1:1–19.

54. Kim M, Baek S. The structural relationship of factors in Korean fifth graders' L2 listening proficiency. 영어영문학21 [English 21]. 2017;30 (3):337–364.

55. Pekrun R, Linnenbrink-Garcia L. *International Handbook of Emotions in Education*. Routledge; 2014.

56. Ringle CM, Wende S, Becker J-M. *SmartPLS 3*. Boenningstedt: SmartPLS; 2015.

57. Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: myths and truths about mediation analysis. *J Consum Res*. 2010;37(2):197–206. doi:10.1086/651257

58. Vinzi VE, Chin WW, Henseler J, Wang H. Handbook of partial least squares: concepts. *Methods Appl*. 2010. doi:10.1007/978-3-642-16345-6

59. Vandergrift L, Baker S. Learner variables in second language listening comprehension: an exploratory path analysis. *Lang Learn*. 2015;65 (2):390–416. doi:10.1111/lang.12105

60. Hwang GJ, Hsu TC, Hsieh YH. Impacts of different smartphone caption/subtitle mechanisms on English listening performance and perceptions of students with different learning styles. *Int J Hum Comput Interact*. 2018;35(4–5):333–344. doi:10.1080/10447318.2018.1543091

61. Lai C, Gu M. Self-regulated out-of-class language learning with technology. *Comput Assist Lang Learn*. 2011;24(4):317–335. doi:10.1080/09588221.2011.568417

62. Wu J, Lee MCL. The relationships between test performance and students' perceptions of learning motivation, test value, and test anxiety in the context of the English benchmark requirement for graduation in Taiwan's universities. *Lang Test Asia*. 2017;7(1). doi:10.1186/s40468-017-0041-4

63. Chang ACS. Second-language listening anxiety before and after a 1-yr. intervention in extensive listening compared with standard foreign language instruction. *Percept Mot Skills*. 2010;110(2):355–365. doi:10.2466/PMS.110.2.355-365