



BIROn - Birkbeck Institutional Research Online

Li, Chengchen and Dewaele, Jean-Marc and Hu, Y. (2023) Foreign language learning boredom: conceptualization and measurement. *Applied Linguistics Review* 14 (2), pp. 223-249. ISSN 1868-6303.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/42611/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact lib-eprints@bbk.ac.uk.

Foreign language learning boredom: conceptualization and measurement¹

Abstract

This article reports on an investigation in two steps conducted among Chinese university non-English-major EFL students and English teachers. In Study 1, 22 students and 11 English teachers were interviewed and 659 students responded to an open questionnaire, recalling and describing their experiences or perceptions of boredom in English learning. A qualitative analysis revealed that more than 90% of students reported boredom in FL settings. The data allowed a multidimensional conceptualization of foreign language boredom (FLB), empirically supporting the control-value theory in educational psychology. Study 2 used the new conceptualisation to develop and validate a new scale for FLB. Through surveying 808 students in a pilot study and 2223 in the main study, a 7-factor *Foreign Language Boredom Scale* with 32 items was identified and validated using exploratory and confirmatory factor analyses as well as further validity and reliability analyses. It is argued that FLB is a crucial addition to the emerging field of FL classroom emotion research.

Key words: Foreign Language Boredom (FLB), Foreign Language Boredom Scale (FLBS), control-value theory (CVT), second language acquisition, emotion

Introduction

Emotion has long been shunned in foreign language (FL) learning - with the exception of anxiety (Dewaele & Li, 2018) - because emotion was considered as inferior to more “objective” cognitive factors (Dewaele, 2019; Prior, 2019). However, there has been a steady increase in the attention to emotion over the past decade (Prior, 2019). A growing number of SLA researchers agree that learners experience diverse emotions in language learning as they do in other educational contexts, which can affect the learning process and the performance (Dewaele & Li, 2020a).

The emotional spectrum has expanded beyond anxiety to include various emotions that FL learners may experience, such as enjoyment, excitement, hope, love, pride, interest, joy, gratitude, shame, guilt, distress, anger and sadness (e.g., Dewaele & MacIntyre, 2014; Li, 2020; MacIntyre et al., 2019; Teimouri, 2018). Foreign language enjoyment (FLE) and anxiety (FLCA) have been at the forefront of this new wave of emotion research with studies focusing on their factor structure, correlation, correlates, antecedents and consequences (e.g., Dewaele, Witney, Saito, & Dewaele, 2018; Khajavy, MacIntyre, & Barabadi, 2018; Li, Dewaele, & Jiang, 2019). The popularity could be largely attributed to the availability of psychometrically sound FL-specific measurements of the two emotions (Teimouri, 2018), namely the *Foreign*

1. Pre-print of Li, C., Dewaele, J.-M. & Hu, Y. (2020) Foreign Language Learning Boredom: Conceptualization and Measurement. *Applied Linguistics Review*
<https://doi.org/10.1515/applirev-2020-0124>

Language Anxiety Classroom Anxiety Scale (Horwitz, Horwitz, & Cope, 1986) and the *Foreign Language Enjoyment Scale* (Dewaele & MacIntyre, 2014). In other words, the paucity of FL-specific measures for assessing other emotions might explain why little research has focused on them so far (Teimouri, 2018).

One of the emotions that deserves closer attention in SLA is boredom which has been well-researched in educational psychology given its multifarious repercussions in learning (Putwain et al., 2018). A number of recent studies have addressed its occurrences in FL contexts and its potential negative effects on FL learning, borrowing instruments from general psychology (e.g., Dewaele & Li, 2020a; Li, 2020; Kruk, 2019; Pawlak, Zawodniak, & Kruk, 2020). The very first instrument measuring FL-specific boredom was *Boredom in Practical English Language Classes Questionnaire (BPELC)* (Kruk & Zawodniak, 2017; Pawlak, Kruk, Zawodniak, & Pasikowski, 2020). The instrument was developed among a small group of 107 English majors in Polish educational context. As reviewed subsequently, the psychometric properties of this instrument remain to be further validated, especially its construct validity and reliability, and in a larger sample of non-English-major EFL learners, in order to avoid inconsistent findings or discrepancies (Goetz, Frenzel, Pekrun & Hall, 2006; Teimouri, 2018). The surge of interest in boredom means that there is an urgent need for a solid, valid and reliable instrument to capture the phenomenon of boredom in FL learning in various contexts, allowing accurate measurement of its occurrence and identification of its sources. Such research would offer FL researchers an operational tool for future research on boredom and it would offer teachers insights into ways to counter it.

Literature review

Academic Boredom

Boredom is an unpleasant emotional or psychological state, associated with low physical arousal and cognitive stimulation, as well as specific time perceptions (e.g., slowing down or standing still) and action tendencies (e.g., withdrawal from ongoing boredom-inducing situations through cognitive or/and behavioral disengagement) (Goetz & Hall, 2014). It permeates educational settings, consistently being detrimental for individual academic functioning, negatively affecting individual behavior, engagement, cognition, interest, curiosity, motivation, learning strategies, performance, and outcomes (Daniels, Tze, & Goetz, 2015; Frenzel, 2014; Götz & Hall, 2014). It is perceived as inconspicuous and silent, which may partially explain why teachers have not paid much attention to it (Macklem, 2015). Also, teachers, although being aware of students' boredom, may attribute it to students' personality variables, such as laziness and trait anxiety or depression and deflect any responsibility about its occurrence (Macklem, 2015).

As the present study is concerned with boredom experienced in FL learning, the following review will be primarily focused on boredom in academic settings, especially its relevant assumptions, theories and findings in educational psychology. More specifically, the literature review will focus on its conceptualization, i.e., the definition of the concept and the description of the indicators to be used in its operationalization and its measurement.

Boredom and Its Multi-dimensionality

According to the control-value theory (CVT) of academic emotions in contemporary educational psychology (Putwain et al., 2018), boredom is a multidimensional construct. Like other discrete academic emotions (e.g., enjoyment and anxiety), boredom can be distinguished by a three-dimension taxonomy: valence, activation/arousal, and objective focus (Pekrun, 2006). *Valence* pertains to the pleasantness/unpleasantness of the emotion, and *activation* concerns physical or/and cognitive activation/deactivation associated with the emotion, and *objective focus* pertains to whether the emotion is activity-related or outcome-related (arising from an ongoing activity or evoked by past or future outcomes of these activities). Following the taxonomy, boredom can be conceptualized as a *negative, deactivating* emotion arising from *ongoing learning activities or tasks*, enjoyment as a *positive, activating* emotion arising from *ongoing learning activities or tasks*, and anxiety as a *negative, activating* emotion evoked by envisioned results related to *future learning outcome or performance* (Pekrun & Perry, 2014).

Boredom and Its Antecedents

The CVT posits that control-value appraisals are two proximal antecedents of boredom, and that other distal antecedents act on these two proximal antecedents to further affect academic emotions. Distal antecedents include personal antecedents, such as gender and achievement goals, and situational antecedents, such as social environment, teacher and peer behaviors (Pekrun, 2006). *Control appraisal* refers to self-perceived control over achievement-related activities and outcomes. *Value appraisal* refers to perceptions of meaning or value of achievement-related activities and outcomes and can be further distinguished as *intrinsic* and *attainment value* appraisals.

According to the CVT, boredom, for example, is assumed to have a ***curvilinear relationship*** with control appraisals and a negative relationship with value appraisal. Specifically, boredom is assumed to be evoked by self-perceived extremely low or high controllability over an ongoing activity or task. In addition, boredom is assumed to be evoked when an ongoing learning activity or task is perceived as irrelevant or meaningless.

These assumptions on the links between control-value appraisals and boredom have been partially supported in different educational contexts, for example, in mathematics (Putwain et al., 2018), and in psychology and education (Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010). Consistent evidence has shown that both control and value appraisals negatively predict boredom. However, it remains unknown whether there is a curvilinear relationship between control appraisals and boredom experience, namely both high control and low control predicting high level of boredom. In addition, these empirical findings remain to be corroborated in the specific field of FL learning and teaching considering the assumption of domain-specificity of academic emotion (Goetz et al., 2006).

Types of Boredom

Trait boredom and state boredom. According to the CVT, boredom, temporally generalized, could be basically grouped into *state boredom* and *trait boredom*. The former refers to momentary, transient boredom experience instigated in reaction to a

given situation, while the latter refers to a relatively stable proneness, disposition or habitual and recurring boredom experience in relation to a learning activity or a subject (Putwain et al., 2018). Although inherently different, it is easy to assume that people who have higher trait boredom are more likely to experience corresponding state boredom (Putwain et al., 2018). This classification addresses boredom from both a situational and an individual perspective.

Researchers have also differentiated between classroom boredom and homework boredom (Goetz et al., 2012). Both are found to be closely linked with achievement outcomes, and they need to be measured and assessed separately (Macklem, 2015). In the present study, a distinction is made between different types of boredom (trait and state) in different sub-contexts of FL learning (class and home).

Measures of Boredom in Psychology and Educational Psychology

There are a number of measurements for boredom in different disciplines. Vodanovich and Watt (2016) reviewed 16 boredom scales measuring different types of boredom, including: 1) trait boredom, 2) context-specific trait boredom, 3) state boredom, and 4) context-specific state boredom.

The *Boredom Proneness Scale* (BPS) (Farmer & Sundberg, 1986) is a well-known measurement for trait boredom. It consists of 28 items assessing individual general tendency to experience boredom. An example item is “Much of the time, I just sit around doing nothing”. It had a true-false response format in its initial form.

An example scale for context-specific trait boredom is the *Free Time Boredom Scale* by Ragheb and Merydith (2001). It includes 33 items to capture individuals’ general tendency to experience boredom in their leisure time. It is responded to on a 5-point Likert scale. An example item is “I do things below my physical ability”.

A well-established scale for state boredom is the *Multidimensional State Boredom Scale* (MSBS) by Fahlman, Mercer-Lynn, Flora and Eastwood (2013). It is comprised of 29 items answered on a 7-point scale. It measures the momentary, transient experience of boredom in reaction to a certain situation in general. Five factors were found, namely *disengagement, high arousal, inattention, low arousal and time perception*. An example item is “I am stuck in a situation that I feel is irrelevant”.

In terms of context-specific state boredom, three widely used scales in educational context are: 1) *the Boredom Subscale* of the *Achievement Emotions Questionnaire* (AEQ) by Pekrun, Goetz, Frenzel, Barchfeld and Perry (2011), 2) the *Academic Boredom Scale* (ABS) by Acee et al.(2010), and 3) the *Precursors to Boredom Scales* (PBS) by Daschmann, Goetz and Stupnisky (2011). They all measure boredom experience in the specific academic contexts. The AEQ assesses several achievement emotions, including boredom that are prevalent in academic settings. The initial version of *Boredom Subscale* consisted of 31 items, while the short version consists of 22 items, with 11 items measuring class-related boredom and learning-related boredom each. The items are responded to on a 5-point Likert scale. An example item for class-related boredom is “I’m tempted to walk out of the lecture because it is so boring”. An example item for learning-related boredom is “While studying I seem to drift off because it’s so boring”. The ABS includes 10 items that are responded to on a 9-point scale. The items distinguish task-focused boredom from self-focused boredom. All the items follow an item stem “In that situation, to what extent did you”. An example item for task-focused boredom is “Find the activity dull?”, and an example item for self-focused boredom is “Become frustrated or

annoyed?”. The PBS is comprised of 22 items assessing eight distinctive precursors of boredom in school settings: *Monotony*, *Lack of Meaning*, *Opportunity Costs*, *Being Over-Challenged*, *Being Under-Challenged*, *Lack of Involvement*, *Teacher Dislike*, and *Generalized Boredom*. The items are arranged on a 5-point Likert scale. Example items are “The subject matter in class has no meaning in my life” for the dimension of Lack of Meaning, “We always do the same thing in class” for the dimension of Monotony, “The subject matter is so easy” for Being Under-Challenged, and “My instructor isn’t likeable” for Teacher Dislike.

Apart from the above three popular scales for boredom in academic context, the *Homework Boredom Scale* (HBS) was designed to capture students’ boredom during homework. HBS is a subscale of the *Homework Emotion Scale* Goetz et al. (2012) developed on the basis of AEQ. It consists of four items arranged on 5-point Likert scale. The item is provided in a form showing readiness to be used in different subjects, for example, “When doing my [DOMAIN] homework I can’t concentrate because I am so bored”.

The above-mentioned scales of boredom shed insights on the conceptualization and measurement of boredom in the specific domain of FL learning. For instance, we should take into consideration the different types of boredom among FL learners (e.g., trait/state boredom), with learner-internal and learner-external factors (e.g., teacher and task) in various sub-contexts (class and homework).

Boredom in FL Learning

Interest in boredom is recent in FL research (Pawlak, Kruk, Zawodniak, & Pasikowski, 2020) and has been mostly conducted in Polish FL contexts (Pawlak, Kruk, & Zawodniak, 2020a), with the exception of Chapman (2013) who adopted a mixed methods approach to investigate the boredom experiences (beliefs, causes, and associated behaviors) of 57 American learners of German as a FL over a period of four weeks. Quantitative data on boredom was gathered by means of the above-mentioned *Boredom Proneness Scale*.

Zawodniak, Kruk and Chumas (2017) analysed diaries of 30 Polish university English majors and found that boredom was pervasive in their English classes. Kruk (2016) sought to examine variation in levels of boredom, anxiety and motivation, experienced by 16 English majors over a semester during their visits to *Second Life*. Quantitative (self-ratings on the levels) and qualitative data were gathered through a three-part session log, consisting of background information section, self-rating section, and a section for comment on boredom, anxiety and motivation experiences. Using a similar research design, Kruk (2019) explored the changes of levels of willingness to communicate, motivation, boredom and anxiety experienced by two Polish advanced learners of English (English majors) during sessions in their *Second Life*. In another study, Pawlak, Kruk and Zawodniak (2020b) investigated three Polish university English majors’ boredom experiences in three 90-minute English classes, exploring: (1) the patterns displayed, (2) the relationship between general proneness to boredom and boredom in English classes, and (3) the trajectories and their factors. The tool to gather quantitative data was the *Boredom in Practical English Language Classes Questionnaire (BPELC)* (Kruk & Zawodniak, 2017). Qualitative data was gathered through short paragraphs participants wrote about their experiences of boredom during each class. Pawlak, Kruk, Zawodniak and Pasikowski (2020) used the *BPELC* to gather data from 107 participants. Using exploratory factor analysis (EFA), they identified two factors underlying the construct of boredom in practical

English classes: 1) *Disengagement, monotony, and receptiveness*, and 2) *Lack of satisfaction and challenge*.

The recent research on FL boredom, initiated by the group of Polish scholars, opens new perspectives and presents fresh challenges. Firstly, there is an urgent need to carry out research outside the Polish educational context. Secondly, the target population needs to extend beyond English majors at university level. Thirdly, most of the previous studies on FL boredom were based on either a purely qualitative approach or a mixed methods approach with very small sample sizes. In order to allow some degree of generalization about boredom experiences of FL learners, a much larger sample size is needed. Fourthly, conceptual and methodological development is what pushes applied linguistics research forward (Phakiti et al., 2018).

Kruk and Zawodniak's (2017) 27-item *BPELC* is the first instrument to measure boredom in FL learning and contributes substantially to emotion research in SLA, drawing scholarly attention to this neglected but prevalent negative emotion. However, its psychometric properties have only just been reported (Pawlak et al., 2020) and the following aspects may need to be addressed. Firstly, the sample size is relatively small for the validation of an instrument, including the EFA conducted in the study. The rule of thumb has been at least ten respondents for each scale item (Boateng et al., 2018). Thus, for the 27-item *BPELC*, the ideal sample size should be larger than 135 participants. However, it is difficult to find a consensus on how large should a sample of subjects be (DeVellis, 2016). It could be only agreed that "the sample should be sufficiently large to eliminate subject variance as a significant concern" (DeVellis, 2016, p. 130) and that "there are several risks in using too few subjects" (p. 130), for example, "the patterns of covariation among the items may not be stable" (p. 130) and "the development sample may not represent the population for which the scale is intended" (p. 130). Secondly, besides EFA, other statistical analyses should be conducted to test other various kinds of psychometric properties for a newly-developed instrument, such as construct validity of the overall scale (through EFA in one sample and confirmatory factor analysis in another sample), convergent validity and discriminant validity for the subscales, and criterion validity in comparison with related well-established scales, as well as reliability (e.g., internal consistency and test-retest reliability) of both the overall scale as well as the subscales (DeVellis, 2016). Thirdly, little is known about the process of the generation of the item pool, nor the sources of the items, discounting the methodological transparency and theoretical foundations. Fourthly, boredom occurring in relation to L2 learning may be trait boredom or/and state boredom, classroom boredom and out-of-class boredom (especially homework boredom). Considerations of these types may provide a more in-depth understanding and measurement of boredom in English learning. Finally, the instrument was developed for a very specific population, namely English majors in a Polish university, and its applicability to non-English majors remains to be examined. According to the CVT, control and value appraisals follow the principle of domain specificity, suggesting that students' academic emotions should also be constructed in subject-specific ways (Pekrun & Perry, 2014). It is not difficult to imagine that an English major may have a distinctive academic self-concept and values around English that differ from non-English majors, which could have repercussions on the profiles, conceptualizations and emotional experiences in English learning.

The literature review has revealed substantial research on boredom in educational psychology and a recent interest in applied linguistics. We argue that the development of a psychologically sound questionnaire would further increase interest in FL

boredom. The present study thus aims to answer the following three research questions:

RQ1: Do learners experience boredom in EFL settings?

RQ2: How do learners conceptualize FLB in an EFL setting?

RQ3: Is the FLBS a psychometrically sound tool for measuring learners' boredom reactions in relation to EFL learning? To what extent is the scale valid, reliable and stable?

Methodology

As we seek to develop and validate a new instrument, we adopted an exploratory sequential design. Its intent is to: (1) first explore a topic through qualitative data collection and analysis, (2) then develop a new measure or instrument based on the results from the qualitative analysis (e.g., themes and indicators), and (3) finally apply the new instrument and test it out (Creswell, 2014).

Correspondingly, we conducted two separate studies. Study 1 addressed the first two RQs. RQ1 was probed through qualitative analysis of interview data obtained from students' responses and their English teachers, of students' responses to open questions. RQ2 was answered by weighing the qualitative findings in light of the literature in educational psychology.

Study 2 addressed RQ3. Based upon the conceptualization and theoretical construct of FLB addressed in RQ2, students' responses obtained in the interviews and open questions of Study 1 were further used to develop an item pool of the FLBS. The psychometric properties were assessed in Study 2 in two different samples in three stages. The psychometric properties assessed included multiple kinds of validity (i.e., face validity, construct validity, convergent validity, discriminant validity and criterion validity) and reliability (i.e., internal consistency and test-retest reliability). In order to conceptualize FLB and develop a psychometrically sound scale for FLB, we designed four phases with different foci (see Table 1). The research design and questionnaire obtained ethical approval at the author's institution.

Table 1 Design and Procedure

Study	Phase	Sample	Instrument	Data analysis	Tool	Time (2019)	
Study 1 Occurrence & conceptualization of boredom	Qualitative Phase	$n_1 = 22$	Interview	Qualitative analysis	Excel	10 May	
		$n_2 = 11$ (Ts)				-	
		Sample 1 ($N1 = 659$)	Open questionnaire			15 July	
Study 2 Development & Validation of the FLBS	Quantitative Phases: Item pool generation		69-item FLBS	Face validity	Professional assessment	20 Oct. - 1 Nov.	
		Initial scale exploration	Sample 2 ($N2 = 808$)	47-item FLBS	EFA reliability validity analyses	SPSS 19.0	20 Oct. - 1 Nov.
		Scale confirmation	Sample 3 ($N3 = 2223$)	32-item FLBS	CFA reliability validity analyses	MPLUS 7 SPSS 19.0	5 Nov. - 15 Nov.

(<i>n</i> = 90)	Formal 32-item FLBS	test-retest reliability	SPSS 19.0	25 Nov. - 15 Dec.
------------------	---------------------------	----------------------------	-----------	----------------------

Note: Ts= Teachers; all the participants were students except the 11 English teachers interviewed in Study 1 indicated as Ts.

Recruiting Strategy and Data Collection

A convenience sampling strategy was adopted in the two sub-studies. The authors contacted their colleagues and associates who also taught college English at universities in China and invited them to recruit their students in the study. They were informed of the main purpose (without knowing exactly the specific research questions) and the approximate duration of the interview as well as the questionnaire survey. They were also ensured of the anonymity and voluntary nature of participation. Once the teachers agreed, they helped to invite the students in their target classes to participate in the interview or the questionnaires in different time points. Before proceeding to the interview or questionnaires, students were provided with information concerning the nature and purpose of the survey, the approximate duration, the anonymity handling of data, and their rights for non-participation as well as withdrawal at any time.

The research group first had face-to-face interviews with 22 students and 11 English teachers. The questionnaires in both Study 1 and Study 2 were uploaded to an online survey tool www.wjx.cn, generating QR codes, which were then forwarded to the participating teachers via social apps (e.g., Wechat and QQ). The teachers were asked to insert the QR codes in their PPT slides and have their students fill in the questionnaires during the breaks between regular classes. To ensure full understanding and expression, all the instruction and responses were provided in Chinese.

Study 1

Study 1 had three main goals: 1) to provide evidence on the occurrence of boredom among Chinese university non-English-major EFL learners from both the learners' perspective and the teachers' perspective, 2) to provide an empirical basis for the conceptualization of the construct of FLB, and 3) to provide reliable information sources for generating the item pool of the questionnaire.

Participants

Student participants in the interview. Twenty-two undergraduate students (14 males and 7 females; average age = 19.05, *SD* = 1.13, range: 18-22) were recruited from five universities in different regions in China that had differential academic rankings (See Table 2 for detailed information). They were interviewed on their experiences of emotions in their English classes with a focus on boredom. Example guideline questions for the semi-structured interview are: 1) *Have you ever experienced any emotions in English learning? If so, when and why?* 2) *Have you ever got bored in English learning? If so, When? Could you give any examples?*

Table 2 Information of student interviewees

Student	University	Gender	Age	Length of English learning (Years)	Grade	Major
S1	A	Male	20	12	Year 3	Computer

S2	B	Female	18	11	Year 1	Psychology
S3	C	Male	18	11	Year 1	Japanese
S4	C	Female	18	10	Year 1	Medicine
S5	B	Male	18	9	Year 1	Math
S6	C	Male	18	8	Year 1	Medicine
S7	C	Male	19	7	Year 1	Physics
S8	C	Female	18	10	Year 1	Medicine
S9	B	Male	19	10	Year 1	Biology
S10	B	Female	19	15	Year 1	History
S11	A	Female	19	13	Year 1	Business
S12	C	Male	19	10	Year 1	Photoelectricity
S13	C	Male	18	10	Year 1	Medicine
S14	A	Female	19	11	Year 1	Management
S15	C	Male	18	9	Year 1	Medicine
S16	D	Female	21	12	Year 1	Literature
S17	D	Male	21	12	Year 1	Radio and television programming
S18	E	Male	22	13	Year 1	Financial Management
S19	E	Male	19	10	Year 1	Mechanics
S20	E	Male	19	7	Year 1	Physics
S21	E	Female	20	17	Year 1	Financial Management
S22	C	Male	19	10	Year 1	Photoelectricity

Note: S= Student

Student participants in the open questionnaire. Six hundred fifty-nine undergraduate students participated in the online open questionnaire and provided their responses. They were enrolled in 15 universities at different academic levels from the north to the south of China. A majority were Year 1 students ($n = 391$, 59.3%), with 264 in Year-2 (40.1%), 2 in Year 3 (0.3%) and 2 in Year 4 (0.3%). There were 230 males (34.9%), 384 females (58.3%), and the rest ($n = 45$, 6.8%) preferred not to say. The mean age was 18.5 ($SD = .95$), with a range of 16 to 29. All were non-English majors from 53 different disciplines, including the humanities (e.g., philosophy, psychology, literature, history, and arts), and the natural sciences (e.g., maths, physics, chemistry, geology, and biology). The questions asked in the open questionnaire were: 1) *How do you translate “boredom” into Chinese?* 2) *Could you describe the feeling of boredom in details?* 3) *Have you ever experienced boredom in relation to English learning? When?*

Teacher participants in the interview. Eleven English teachers (Gender: 6 females, 5 males; Age: $M = 41.73$, $SD = 8.67$, range: 25-55; years of teaching: $M = 22.10$, $SD = 11.01$, range: 2-47) from four universities with different academic rankings were also interviewed (See Table 3 for detailed information). The questions probed into their perceptions of their students’ emotional experiences in English, boredom in particular. Example guideline questions are: 1) *Have you ever perceived students’ emotions in class? If so, when and why?* 2) *Do you think students may get bored in class? When and why? Could you give any examples?*

The language used in the interview as well as in the questionnaires in the present study was Chinese, with a view to ensuring full understanding as well as understanding.

Table 3 Information of teacher interviewees

Teacher	University	Gender	Age	Length of English instruction	Weekly teaching hours
T1	C	Female	43	18	16
T2	B	Female	47	25	12
T3	C	Male	44	22	16
T4	B	Female	41	19	12
T5	A	Male	37	15	15
T6	C	Female	29	2	12
T7	A	Female	54	30	15
T8	A	Female	25	47	15
T9	C	Male	55	32	16
T10	C	Female	42	16	12
T11	D	Female	42	17	12

Results

Boredom in the EFL Class

An analysis of responses in the interview and open questionnaire involved identifying the excerpts about boredom experiences and extrapolating their typical occurrence in class. The 22 students in the interview reported 49 instances of boredom they experienced in English class. One of them recalled six boredom-inducing situations. That is, on average, they each reported 2.33 boredom-inducing events or situations. In the open questionnaire, 610 out of the 659 (92.6%) student participants also recalled boredom-inducing situations either in FL class or after class (homework) in relation to English learning. A large majority of students (92.8%: $(610+21)/(659+21)$) thus reported experiencing occasional boredom episodes in FL settings.

The 11 English teachers interviewed reported that students did occasionally get bored in class. Each of them recalled at least one episode when students got bored. In total, they reported 34 instances. This indicated that they each reported an average of three episodes, events or situations that led to their students' boredom.

Teachers and students either explicitly used the word of boredom/bored/boring/bore, or closely related words, such as dull, uninteresting, monotonous, lack of interest or meaning, tedious, restless, and dreary. These results suggest that episodes of boredom occur in FL learning settings, in and out of class.

Conceptualization of Foreign Language Boredom

Three dimensions and two appraisal antecedents. A total of 384 (58.3%) out of the 659 participants in the open questionnaire provided their responses to the question of “*Could you describe the feeling of boredom in detail?*”. Their descriptions were summarized and displayed in Table 4. Boredom experienced in English learning could be conceptualized along three dimensions with two antecedents. The three dimensions were 1) valence (positive or negative), 2) arousal (degree of activation), and 3) object focus (activity-related or outcome-related). The two appraisal antecedents were 4) control (degree of controllability), and 5) value (relevance/meaning/interest) appraisals. According to Table 4, Foreign Language Boredom is a *negative* emotion with extremely low degree of *activation/arousal* that arises from *ongoing activities* (example indicative words are “disengagement”, “inattention”, “time dragging”).

These activities are typically *over-challenging* or *under-challenging* and/or of *little significance, relevance, or meaning* to the learners. More specifically, 1) FLB was strongly characterized by *negative feelings* revolving around these situations or activities, such as uninterest, restlessness, anxiety, frustration, helplessness, dislike, unpleasant state of passiveness, sadness, impatience, depression, blankness, guilt, tiredness, struggling sentiment of time dragging, sleepiness and dissatisfaction. 2) Boredom was a *low-arousal emotion* in that it was strongly linked to inactivity, disengagement, inattention, mind wandering and loss of goal in ongoing activities, as well as various academic withdrawal, escaping, or avoidance behaviors. 3) The activity-related object focus was indicated by academic withdrawal from or disengagement in an *ongoing activity*, which is perceived as uninteresting, meaningless, and not sufficiently stimulating. 4) & 5) Boredom occurred in *over-challenging* or *under-challenging* (repetition, monotony, lack of diversity, variability and novelty) tasks or activities, especially those which were perceived by learners as *unimportant, meaningless, and irrelevant*.

Table 4 Words (expressions) used to describe the feeling of boredom and the frequency of its occurrences in the data.

words describing boredom	N	words describing boredom	N	words describing boredom	N
lack of interest (no interest)	76	desire to escape (withdraw, end, disengagement)	11	desire of self-distraction	2
restlessness (agitation, annoying)	66	Frustration	11	guilt	1
lack of meaning	63	dislike	10	depression	1
low arousal (down, inactivity, not stimulating, no desire to do anything)	31	mind wandering	9	dissatisfaction (lack of a sense of achievement)	1
loss of goal (not knowing what to do and why)	29	tiredness (physical or/and mental tiredness, exhaustion)	9	lack of discipline	1
disengagement	23	difficulty (over-/under challenged)	7	negative feeling	1
unhappiness (unpleasantness)	21	blankness	5	no enthusiasm	1
withdrawal of interest	20	helplessness	5	no passion	1
dull	17	no motivation	4	sadness	1
sleepiness	15	time dragging	3	sigh	1
inattention	13	trait-boredom	2	struggling	1
unpleasant state of passiveness	13	bewildering thoughts	2	self-disclosure	1
anxiety	12	impatience			

N: frequency in the reports.

Two types. In addition to the three dimensions and the causal links, two types of boredom in FL learning were identified in students' reports in the open questionnaire as well as in the interview. The first type was the momentary *state boredom* evoked in reaction to a given FL learning situation/activity and the second type was the more habitual, recurrent and stable *trait boredom* in relation to FL learning. The following excerpts exemplified state boredom as well as trait boredom in the FL learning context.

S1 (male, 20). *I felt so helpless and bored when I was asked to do an overwhelmingly difficult classroom exercise that I could do nothing but to keep myself seated in the classroom, turning over the textbook.* (student interview extract)

1)

S3 (male, 18). When the English teacher kept repeating things that I had already **had a good mastery of**, I felt bored and wouldn't follow her lecturing. (student interview extract # 2)

The above extracts illustrated two typical boredom-inducing situations in the English class. The boredom here was a momentary, here-and-now emotional experience instigated in reaction to over-challenging and under-challenging tasks in class at a short timescale. It is thus transient and fleeting. In other words, it is subject to and highly dependent on the situational or contextual dynamics. For example, if the teachers became aware of her/his repetition in class, and made appropriate adjustment such as engaging students in a task of novelty, emotion other than boredom may arise in response.

S1 (male, 20). I don't *like the subject of English*. *It is so difficult and there is a huge imbalance between the efforts and outcomes of learning, which makes so bored and frustrated*. However, I have to learn it because... (student interview extract # 3)

S21 (female, 20). *Honestly speaking, I think English is of no relevance to our major. I feel so bored and sleepy in English class, especially bored of English words and grammar...I think the main reason is I don't have any interest in English and I don't care about English at all*. (student interview extract # 4)

S16 (female, 21). *The moment I think of English class, I feel bored...* (student open question extract # 1)

The above extracts suggest that boredom is a relatively stable and individualized emotional disposition or proneness in FL learning over a long timescale (e.g., course-scale, term-scale), being habitually and generally bored with the subject of English and the English class.

Summary

The qualitative analysis of teachers' and students' reports in the interview as well as in the open questionnaire revealed the occurrences of EFL learners' boredom. Students' descriptions of the feeling and experience of boredom were categorized along three dimensions (object focus, valence, arousal) with control-value appraisals as antecedents and as two types (trait boredom and state boredom).

Study 2

The results of Study 1 on the occurrences and conceptualization of FLB warrant further exploration. Similar to previous research on foreign language classroom anxiety (Horwitz et al., 1986), enjoyment (Dewaele & MacIntyre, 2014), shame and guilt (Teimouri, 2018), an FL-specific and psychometrically sound measurement is crucial for further research on FLB. Thus, the following section is devoted to the development and validation of the FLBS.

Developmental Stage

The first phase of Study 2 was conducted to develop the initial FLBS.

Factors proposal: Before item generation, we proposed a 7-factor structure for FLB based upon the synthesis of 1) the relevant literature on boredom in education psychology and SLA (e.g., Farmer & Sundberg, 1986; Fahlman et al., 2013; Acee et

al., 2010; Goetz et al., 2012; Kruk, 2016, 2019), and 2) the responses of students and teachers in the interview and questionnaire survey in Study 1. The seven proposed factors were 1) *Foreign Language Class Boredom*, 2) *Teacher dislike-related Boredom*, 3) *Peer-related Boredom*, 4) *Task/content-related Boredom*, 5) *Material/resource/tool-related Boredom*, 6) *Homework-related Boredom* and 7) *General learning boredom*.

Item pool generation. All the responses obtained from students and teachers in Study 1 were used as sources for the generation of the item pool of the scale. For example, in Study 1, students responded to the open questions: “1) *What do you translate “boredom” into in Chinese?*” 2) *Could you describe the feeling of boredom in details?* The Chinese translation and description (e.g., expressions in Table 4) occurring frequently were chosen as the indicators of boredom in the items of the scale to address the issue of translation inequivalence. Besides, the findings in Study 1 also provided a basis for the generation. For example, the different dimensionality, types (trait and state boredom), and situations (classroom and homework) for boredom identified in Study 1 were addressed in the scale development.

As a result, an initial pool of 69 items was developed based on the following guidelines: The items should: 1) measure boredom experience, proneness or tendencies in relation to FL learning, 2) cover different aspects of FL learning (e.g., settings: in and out of class, learner-internal and learner-external factors: learner, teacher, peers, technology, materials, etc., process: input and output), 3) be likely to be experienced by FL learners. 4) In addition, there should be at least three items measuring each of the seven proposed factors of FLB.

Face validity. The initial scale was assessed by one emotion psychologist and three applied linguists. They were asked to 1) judge whether each item is measuring boredom in the FL context, 2) judge whether each item is well expressed, 3) to choose among the factors that an item belongs to. Disagreement was settled after further discussion, followed by corresponding revisions. Finally, an initial FLBS was obtained. It was arranged on a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

Exploration of the Factor Structure

Phase 2 was conducted to explore and identify the factor structure of the FLBS.

Participants. Sample 2 participated in the first wave of the questionnaire survey. A total of 907 undergraduate students filled in the online questionnaire, and the data of 99 inadequately attentive respondents were removed because their failure rate in trap questions was more than 80% (Liu & Wronski, 2018). Thus, Sample 2 consisted of 808 students. They were from nine universities at differential ranking level in different parts of China. Over half of participants were in Year 1 ($n = 465, 57.5\%$), 339 in Year-2 (41.9%), 2 in Year 3 (0.2%) and 2 in Year 4 (0.2%). There were 252 males (31.2%), 498 females (61.6%), and the rest 58 (7.2%) preferred not to say. The mean age was 18.6 ($SD = .95$), with a range of 16 to 29. They were all non-English majors, and from more than 60 disciplines, 436 being from disciplines in the humanities and 361 from the natural sciences. Eleven students failed to provide that information.

Item analysis

Item discrimination was then carried out. Two groups were created according to their total scores in the 69-item scale: the upper and lower 27% groups. An independent-samples *t*-test was conducted between the two groups for each of the 69 items. No significant difference was found in the scores of nine items (Items 3, 8, 22, 23, 31, 38, 39, 41, 43; all at $p < .01$ level) between the two groups. The nine negatively discriminating items were thus eliminated. The item-total correlation test was then conducted between the score of each item and the overall assessment score. A correlation value smaller than 0.3 was undesirable (Field, 2013) and the corresponding item was dropped. Thus, Item 29 was dropped ($r = .094$).

Exploratory factor analysis

In order to explore the dimensionality of the FLBS and the factor structure of the construct of FLB, and to reduce items, an exploratory factor analysis (EFA) was performed using SPSS 19.0. The Kaiser-Meyer-Olkin (KMO) ($KMO = .956 > .70$) and Bartlett's test ($p < .001$) results were desirable. We then adopted the extraction method of Principal Component Analysis and the rotation method of Varimax with Kaiser Normalization without the seven-factor extraction limits. Following the suggestions of Chen (2016) and Hair et al. (2010) we deleted:

- 1) Items with factor loadings smaller than .40;
- 2) Item that had substantial cross-loadings greater than .40;
- 3) Items that were logically or theoretically inconsistent with other items under the same dimensions/factors.

Table 5 Results of Exploratory Factor Analysis: Factor, item and loading ($N = 808$)

Fac 1	Fac 2	Fac 3	Fac 4	Fac 5	Fac 6	Fac 7	
B5	.820	B48 .745	B58 .717	B62 .702	B21 .726	B67 .849	B40 .584
B6	.801	B47 .740	B56 .702	B61 .689	B26 .707	B66 .830	B27 .550
B4	.789	B45 .740	B55 .671	B60 .649	B20 .678	B68 .788	B28 .486
B9	.764	B46 .731	B57 .659	B63 .587	B25 .665	B69 .730	B42 .485
B10	.739	B44 .711	B30 .472	B59 .491		B65 .666	B33 .469
B12	.711	B49 .702	B34 .461	B64 .436			
B1	.705	B50 .531					
B15	.701	B36 .524					
B2	.686	B37 .520					
B11	.665						
B7	.659						
B13	.655						
Total	27.32	3.33	2.90	2.10	1.76	1.40	1.22
% of Variance	45.54	5.56	4.83	3.49	2.92	2.34	2.04

Note: Factor loadings smaller than .40 were not included.

During this process, 12 items (Items 14, 16, 17, 18, 19, 24, 32, 35, 51, 52, 53, 54) were discarded, which resulted in a 47-item FLBS (Table 5). As proposed in the development stage, seven factors were identified. However, three of them were slightly different from those proposed in the developmental stage. The seven factors identified were named as follows: 1) *Foreign Language Class Boredom*, 2) *Under-challenging Task Boredom*, 3) *PowerPoint Presentation Boredom*, 4) *Homework Boredom*, 5) *Teacher-dislike Boredom*, 6) *General Learning Trait Boredom*, and 7) *Over-challenging or Meaningless Task Boredom*.

Validation of the Psychometric Properties

Phase 3 was conducted to further refine and confirm the factor structure of the 47-item FLBS obtained in the previous EFA, and assess its psychometric properties. More specifically, we assessed its construct validity using confirmatory factor analysis, convergent validity and discriminant validity, criterion validity, reliability and stability.

Participants. Sample 3 was another group of undergraduate students recruited in the second wave of the questionnaire survey. A total of 2223 out of 2393 (92.9%) respondents provided complete and valid data. They were from 12 universities. There were 1940 (87.3%) participants in Year 1, 265 (11.92%) in Year 2, 12 (0.5%) in Year 3, and 6 (0.3%) in Year 4. They were all non-English majors. Slightly over half of the participants were from natural sciences ($n = 1157$, 52.1%), close to one-third were from social sciences and humanities ($n = 683$, 30.7%). The remaining participants ($n = 383$, 17.2%) did not provide that information. Close to half of the participants were females ($n = 1089$, 49.0%), 901 participants (40.5%) were males and 233 participants (10.5%) preferred not to say. Their mean age was 18.3 ($SD=.95$) with a range of 16 to 20.

A group of 118 participants from Sample 3 further participated in the final wave of questionnaire survey (namely the retest). The data of 90 respondents in the second wave and the final wave were valid and successfully matched. There were 40 (44.4%) male students, 37 (41.1%) female students, and 13 (14.5%) preferred not to say. Their mean age was 18.1 ($SD =.68$). They were all in their Year 1 and 8 of them (8.9%) were from subdisciplines in social sciences and humanities, while the rest 82 (91.11%) were from subdisciplines in natural sciences.

Validity tests

Construct validity. To further confirm the factor structure of the FLBS, confirmatory factor analysis (CFA) was conducted by using the data from Sample 3. The CFA was conducted using Mplus 7 with six indices used to assess model fit: (1) χ^2 : model chi-square; (2) df: degree of freedom; (3) CFI: the comparative fit index; (4) TLI: the Tucker Lewis Index; (5) SRMR: standardized root mean square residual; (6) RMSEA: the root mean square error of approximation (RMSEA). According to the benchmarks for acceptable model fit (Kline, 2016), the first round of CFA results (see Table 5) suggested that the seven-factor model with 47 items was not satisfactorily supported.

We then did further CFAs by removing the items of low factor loadings (lower than 0.5) or/and high modification indices (higher than 40, from higher to lower), keeping it in mind that there should be at least three items for each factor (Hair et al., 2010). We finally removed 15 items (Items 9, 10, 11, 12, 30, 34, 36, 37, 40, 42, 44, 50, 58, 59, 64) from the 47-items pool and obtained a 32-item scale of the same seven factors (See Figure 1). The CFA results indicated that the 32-item version was well supported (See Table 6 and Figure 1).

Table 6 Results of Confirmatory Factor Analysis ($N = 2223$)

	χ^2	df	CFI	TLI	SRMR	RMSEA [90% C.I.]
benchmark			> .90	> .90	< .08	< .08
47-item	10023.384***	1013	.884	.876	.884	.063 [.062, .064]
32-item	3950.297***	443	.930	.921	.045	.060 [.058, .61]

Note: *** means significance at < .0001 level; CFI = Comparative Fit Index; TLI= Tucker–Lewis Index; SRMR=Standardised Root Mean Square; RMSEA = Root Mean Square Error of Approximation

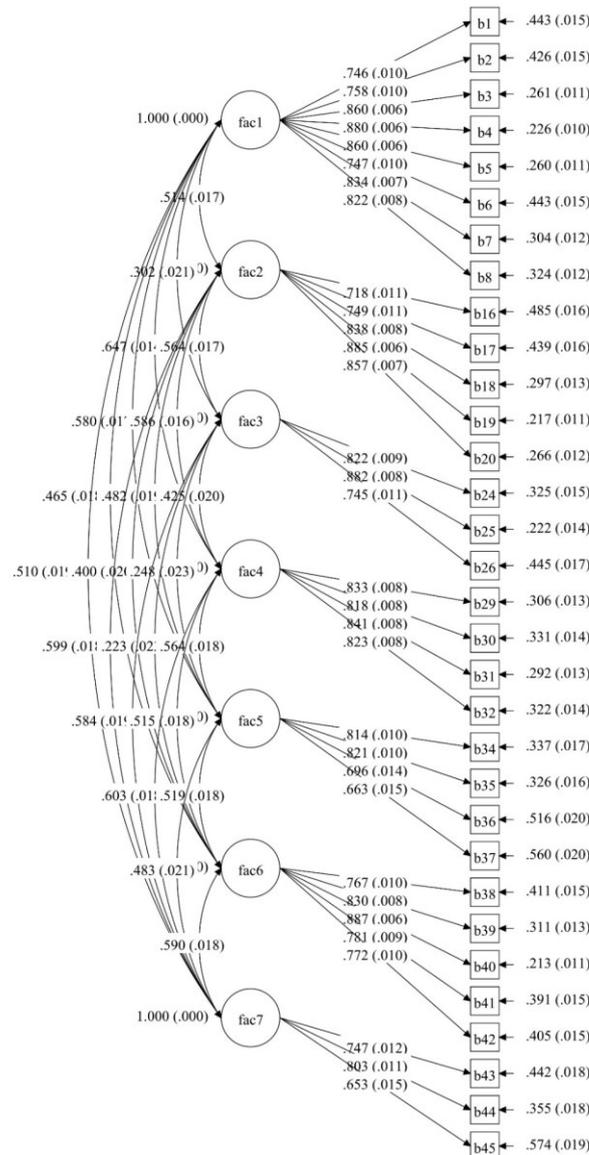


Figure 1. The seven-factor model of Foreign Language Boredom

Normality tests for each item were conducted following the acceptable limits of “± 2” (Gravetter et al., 2020). The Skewness and Kurtosis of all the 32 variables in the scale were within the range of -1.071 to .048 and of -.001 to 1.131 respectively, indicating that the score of each item showed normal distribution.

Convergent validity. Besides the construct validity of the overall FLBS, the validity of each subscale was also testified by assessing the convergent validity. As

displayed in Table 7, the model fit for each subscale was ideal.

Table 7 Convergent validity and model fit indicators of each subscale ($N = 2223$)

Factor	Item	Convergent Validity			Model Fit Indicators						
		λ	P	COM	AVE	CR	χ^2/df	p	CFI	RMSEA	SRMR
Factor 1	1	.746	.000	.557							
	2	.758	.000	.575							
	3	.860	.000	.740							
	4	.880	.000	.774	.664	.940	726.989/20	.000	.950	.012	.030
	5	.860	.000	.740							
	6	.747	.000	.558							
	7	.834	.000	.696							
	8	.822	.000	.676							
Factor 2	16	.718	.000	.516							
	17	.749	.000	.561							
	18	.838	.000	.702	.659	.906	365.180/5	.000	.951	.018	.037
	19	.885	.000	.783							
	20	.857	.000	.734							
Factor 3	24	.822	.000	.676							
	25	.882	.000	.778	.723	.886	.000/0	.000	1	.000	.000
	26	.845	.000	.714							
Factor 4	29	.833	.000	.694							
	30	.818	.000	.669	.687	.898	41.739/2	.000	.993	.095	.011
	31	.841	.000	.707							
	32	.823	.000	.677							
Factor 5	34	.814	.000	.663							
	35	.821	.000	.674	.565	.838	.000/0	.000	1	.000	.000
	36	.696	.000	.484							
	37	.663	.000	.440							
Factor 6	38	.767	.000	.588							
	39	.830	.000	.689							
	40	.887	.000	.787	.654	.904	77.563/5	.000	.989	.081	.015
	41	.781	.000	.610							
	42	.772	.000	.596							
Factor 7	43	.747	.000	.558							
	44	.803	.000	.645	.543	.780	.000/0	.000	1	.000	.000
	45	.653	.000	.426							

Note: COM = communalities; CR = composite reliability; AVE = Average Variance Extracted

Discriminant validity. After assessing the validity of each subscale, we assessed the discriminant validity between each two subscales by comparing the Average Variance Extracted method and r^2 (determination coefficient). As shown in Table 8 and Figure 1, all the AVEs for the seven subscales were higher than their corresponding r^2 , indicating that each subscale showed high discriminant validity.

Table 8 Discriminant validity of each subscale ($N = 2223$)

Subscale	r^2							AVE
	1	2	3	4	5	6	7	
1. Factor1	—							.664
2. Factor2	.504**	—						.659
3. Factor3	.283**	.511**	—					.723
4. Factor4	.602**	.543**	.389**	—				.687
5. Factor5	.504**	.444**	.234**	.470**	—			.565
6. Factor6	.441**	.390**	.210**	.479**	.458**	—		.654
7. Factor7	.455**	.514**	.499**	.522**	.401**	.509**	—	.543

Criterion validity. The criterion validity was assessed by relating the FLBS to previous relevant scales. High correlation was found between the FLBS and the

Achievement Emotions Questionnaire (AEQ)—Boredom Subscale (Pekrun et al., 2011) ($r = .793, p < .001$) and between the FLBS and *the Boredom Susceptibility Scale* (ZBS) (Zuckerman, 1979) ($r = .415, p < .001$), showing strong external validity.

Reliability tests

We assessed two kinds of reliability. Firstly, we assessed the internal consistency (Cronbach's Alpha) of the global scale as well as its seven subscales using the data of Sample 3. The Cronbach's Alphas for the global FLBS and its seven subscales were .949, .940, .905, .852, .898, .841, .900 and .771, indicating high reliability.

In addition, we assessed the test-retest reliability by using the data of the 90 participants selected from Sample 3. The 2-week test-retest reliability for the global scale as well as each subscale were .957, .911, .954, .907, .883, .879, .804, and .863 respectively, indicating that the FLBS had a high level of stability.

Summary

Based upon the responses in Study 1, in Study 2, the FLBS was developed to measure FL learning-specific boredom. A 32-item FLBS with seven factors was identified and confirmed. A series of validity and reliability tests provided strong evidence for the sound psychometric properties of the questionnaire.

Discussion

The first RQ focused on the occurrences of boredom among FL learners during FL learning. The results from different sources (teachers and students) converged to provide robust evidence for the existence of boredom in FL learning, confirming previous empirical findings about undergraduates in German, Canadian and Polish universities (Pekrun et al., 2010; Putwain, Becker, et al., 2018; Zawodniak et al., 2017). As claimed by Putwain et al. (2018), boredom is one of the most prevalently experienced academic emotions negatively affecting learning performance and outcome. Further research into boredom in FL learning (e.g., identifying its causes, considering the various ramifications and associated behaviors, and observing learners' coping strategies) is thus greatly needed (Pekrun et al., 2010).

In response to the second research question, we thus developed a robust conceptualization of "Foreign Language Boredom". The empirical findings obtained from the responses of 659 participants in the open questionnaire substantiated central assumptions of the CVT of achievement emotions in the specific FL learning context. FLB was conceptualized in terms of: (1) the three-taxonomy, (2) causal structure and (3) two different types. Firstly, in accordance with the three-taxonomy of achievement emotions in the CVT (Pekrun, 2006), FLB reported by FL learners was found to be a multidimensional construct, characterized by negative valence, low arousal and being achievement-related activity-focused. Theoretically, this multidimensionality points to the flawed positive-negative dichotomy of emotions and suggests the need for a more nuanced understanding of boredom as well as other academic emotions in FL settings, echoing what Han and Hyland (2019) argued about an FL writing context. The multidimensional approach to conceptualize boredom provides a clear picture of what

boredom is (and is not) and how it differs from other emotions. For example, both boredom and enjoyment arise from ongoing activities, with the former being negative with a low arousal and the latter being positive with a high activation. Both boredom and anxiety are negative emotions, and the former is an activity-related emotion with a low arousal while anxiety is an outcome-related emotion with a high arousal. This suggests that although empirically, boredom was found to be negatively related to some positive emotions and positively related to some negative emotions, such as enjoyment and anxiety respectively (Dewaele & Li, 2020b; Putwain et al., 2018), it does not simply represent a lack of positive emotions, or an overlap with other negative emotions (Götz, & Hall, 2014). Pedagogically, these multidimensional distinctions may help FL teachers better understand the nature of academic emotions, their associated behaviors or contexts, and thus help them better recognize student emotions in class.

Secondly, concerning the antecedents of FLB, the patterns suggest that boredom arises when an ongoing task or activity is perceived as over-challenging or under-challenging, irrelevant, meaningless or not interesting. These are in line with the key assumption of the CVT in that it addresses the curvilinear relationship between control appraisals and boredom experiences as well as the negative relationship between value (intrinsic and extrinsic value) appraisals and boredom. Unlike previous studies using a quantitative approach, the present study extends previous findings on the linear and negative relationship between control appraisals and boredom experiences obtained in mathematics (Putwain et al., 2018). Considering that there are three different types of control-value appraisals, future research could explore whether and how the three types act on boredom experiences, independently, jointly or both as Putwain et al. (2018) did. Future research could also consider other distal antecedents (e.g., gender, achievement goals) and how they act on the two proximal antecedents and then on boredom as assumed by the CVT (Pekrun, 2006).

In addition, students' reports of their boredom experiences clearly distinguished two different types of boredom, namely trait boredom and state boredom. More specifically, FL boredom can be conceptualized as both a context-specific emotion and a trait-like emotion. That is, FL boredom refers to not only the momentary, transient here-and-now emotional experience emerging within classroom or in FL learning at a short timescale, but also a relatively stable and individualized emotional experience in FL learning at a long timescale (e.g., course-scale, term-scale), diachronically and accumulatively shaped by those momentary emotions and personal psychological elements (e.g., belief, identity, and personality traits). This distinction echoes the findings in different contexts (Putwain et al., 2018). Overall, the findings aggregate to provide a sound understanding of the nature and characteristics of FL boredom as well as its conceptual structure, providing a strong basis for the subsequent scale development in FL-specific context.

Based upon the findings in Study 1, the third research question in Study 2 addressed the development and validation of the FLBS. A 32-item FLBS with seven factors was confirmed by showing strong psychometric properties in terms of its internal and external validity, reliability and stability, which may be largely attributed to the sound process of item pool generation based on theoretical assumptions, empirical findings as well as exploratory data in the present study. The strengths of the new instrument are as follows. First, it has an appropriate number of items measuring seven factors indicative of both *trait boredom* and *state boredom* in different learning (sub) contexts (e.g., general English learning, English class, and

homework). As such it can provide a better understanding of the nature of boredom and offer useful pedagogical implications. Secondly, we tested the validity and reliability not only for the global 32-item scale but also for its seven subscales, suggesting that each of the subscales could also be used in future research. For example, Subscale 1- *Foreign Language Class Boredom* and *Subscale 4-Homework Boredom* could be used to measure general boredom proneness in relevance to FL class and FL homework boredom respectively. This is especially meaningful when a distinction between different types of FLB is needed. Thirdly, both learner-internal and learner-external aspects (e.g., the English teacher, learning tasks, peers, and learning materials) are considered in the items. This echoes with the individual-social view of emotion: Emotion is not only a personal subjective experience, but also a socially construed construct (Dewaele et al. 2018; Dewaele & Li, 2020b; Li et al., 2018). Thus, this scale includes the social dimension of boredom and its important role in classroom engagement. Finally, the FLBS is designed for non-English-major EFL learners, distinguishing it again from the very first boredom scale in an FL context, Pawlak et al.'s (2020) *BPELC*.

Thus, the study successfully offers an FL-specific measurement with sound psychometrical properties to capture different types of FL-related boredom (trait boredom and state boredom, class boredom and homework boredom) encountered by non-English-major EFL learners in reaction to individual-contextual factors.

Implications, Future Research Directions, and Limitations

The implications of the present study are theoretical, academic, and pedagogical. The existence of boredom among a majority of Chinese EFL learners is an original finding because it is largely unacknowledged in the teaching profession. Our FLB conceptualization and an associated FLBS may contribute to a better awareness and understanding of the nature, features, dimensionality, causal structure, and different types of “Foreign Language Boredom”. The findings empirically support the main assumptions of the CVT of achievement emotions, including the three-taxonomy, control-value appraisal antecedents, and two different types of achievement emotions. This suggests that the CVT is not only applicable in FL contexts, but can also inspire future research in applied linguistics (e.g., Dewaele & Li, 2020a; Li, 2018). A CVT approach could be adopted to investigate different achievement emotions in FL contexts (Dewaele & Li, 2020a), for example, probing their conceptualizations, antecedents and outcomes, and the reciprocity between them based on the main assumptions of the CVT (Pekrun & Stephens, 2010; Pekrun et al., 2010).

This deeper understanding of FLB is also pedagogically important in that it is a prerequisite for teachers' own practice in handling students' boredom in various situations. For example, FL educators may be inspired by the causal links between control-value appraisals and boredom, and devise interventions to reduce students' boredom experiences by intervening on their perceived control and value of ongoing learning activities or tasks. More specifically, FL teachers could try to design learning activities that are neither under-challenging nor over-challenging to students. Activities work better when they are interesting and of great relevance or value to students (Dewaele & MacIntyre, 2014). Another potentially effective way to alleviate boredom is to boost interest by asking students to reflect on the value of being engaged in a certain kind of learning activity and its relevance to their life (Eastwood

et al., 2007; Hulleman & Harackiewicz, 2009). It is a considerable challenge for teachers to understand and cope with students' emotions, and overcoming this challenge is not only central to establishing an engaging and healthy classroom climate, and building student-teacher rapport, but also pivotal to teachers' own professional development (Chen, 2016) as well as their well-being considering the positive emotional contagion between the teacher and students (Dewaele et al., 2018; Dewaele & Li, 2020b; Gkonou, Dewaele & King, 2020).

Moreover, the FLBS developed in Study 2 exhibited sound psychometric properties, indicating that it could be used to measure boredom of FL learners in future research. To our best knowledge, it is the first measurement of Foreign Language Boredom of non-English-major EFL students. It has the potential to open a new fruitful area of emotion research in FL learning and teaching, providing researchers around the world with a valid and reliable FLBS that could be adapted to the local context. This means adapting items referring to social-cultural factors of teaching and learning practices (Li et al., 2018; Teimouri, 2018).

The paper is not without limitations. Firstly, although we attempted to include as many items related to different boredom-inducing situations as possible based on related theoretical assumptions, empirical findings, as well as interview data in the present study, we did not consider boredom evoked in mobile learning or other forms of extramural learning. Considering the increased use of technology in FL learning, technology-related boredom could be a subject for future research. Secondly, the qualitative data collected at one point in time only provides a preliminary understanding of the relationship between control-value appraisals and FLB. More sophisticated research designs (e.g., a mixed methods approach with a longitudinal research design) are needed to explore how control, intrinsic value and attainment value appraisals shape FLB.

Conclusion

In this study, the occurrence of boredom in Chinese EFL learning was explored qualitatively using data from different sources. A multidimensional conceptualization of FLB was then developed, supporting the central assumptions of the CVT. Based on the qualitative data and conceptualization, the FLBS was developed. It was validated to be psychometrically sound in different large-scale samples. Theoretically, the study introduces an emotion from a neighboring discipline that has only just started to attract the attention of applied linguists. Methodologically, the study provides an operational tool that can be deployed or adapted in future studies. Pedagogically, the study provides FL teachers with a better understanding of students' boredom, enabling them to neutralize it. Undoubtedly, an enjoyable class where there is no time for boredom is crucial for linguistic progress, but it is also beneficial to learners' as well as teachers' well-being.

References

- Acee, T. W., Kim, H., Kim, H. J., Kim, J.-I., Chu, H.-N. R., Kim, M., Cho, Y., Wicker, F. W., & Group, B. R. (2010). Academic boredom in under- and over-challenging situations. *Contemporary Educational Psychology, 35*(1), 17–27.
- Boateng, G. O. , Neilands, T. B. , Frongillo, E. A. , Melgar-Quinonez, H. R. , &

- Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, 6, doi: 10.3389/fpubh.2018.00149.
- Chapman, K. E. (2013). *Boredom in the German foreign language classroom*. Unpublished doctoral dissertation, University of Wisconsin-Madison, Madison, WI, USA.
- Chen, J. (2016). Understanding teacher emotions: The development of a teacher emotion inventory. *Teaching and Teacher Education*, 55, 68-77.
- Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage.
- Daniels, L. M., Tze, V. M. C., & Goetz, T. (2015). Examining boredom: Different causes for different coping profiles. *Learning and Individual Differences*, 37, 255–261.
- Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2011). Testing the predictors of boredom at school: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology*, 81(3), 421–440.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). London: Sage.
- Dewaele, J.-M. (2019). When elephants fly: The lift-off of emotion research in applied linguistics. *The Modern Language Journal*, 103(2), 533–536.
- Dewaele, J.-M., & Li, C. (2018). Editorial: Special Issue Emotions in SLA. *Studies in Second Language Learning and Teaching*, 8(1), 15–20.
- Dewaele, J.-M., & Li, C. (2020a). Emotions in second language acquisition: A critical review and research agenda. In C. Li (ed.), *A Positive Psychology Perspective on Emotions in SLA* [Special Issue]. *Foreign Language World*, 196(1), 34-49.
- Dewaele, J.-M., & Li, C. (2020b). Students' perceptions of teacher enthusiasm and their L2 classroom engagement: The mediating effects of student enjoyment and boredom. Unpublished manuscript.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4, 237–274.
- Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety : The effect of teacher and learner variables. *Language Teaching Research*, 22(6), 676-697.
- Eastwood, J. D., Cavaliere, C. , Fahlman, S. A. , & Eastwood, A. E. . (2007). A desire for desires: boredom and its relation to alexithymia. *Personality & Individual Differences*, 42(6), 1035-1045.
- Elahi Shirvan, M., & Taherian, T. (2018). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: latent growth curve modeling. *International Journal of Bilingual Education and Bilingualism*.
<https://doi.org/10.1080/13670050.2018.1441804>
- Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B., & Eastwood, J. D. (2013). Development and Validation of the Multidimensional State Boredom Scale. *Assessment*, 20(1), 68–85.
- Farmer, R., & Sundberg, N. D. (1986). Boredom Proneness-The Development and Correlates of a New Scale. *Journal of Personality Assessment*, 50(1), 4–17.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Los Angeles and London: Sage.
- Frenzel, A. C. (2014). Teacher emotions. In R. Pekrun & L. Linnenbrink-Garcia

- (Eds.), *International Handbook of Emotions in Education* (pp. 494–519). New York/London: Routledge/Taylor & Francis.
- Gkonou, C., Dewaele, J.-M., & King, J. (Eds.) (2020). *The Emotional Rollercoaster of Language Teaching*. Bristol: Multilingual Matters.
- Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006). The domain specificity of academic emotional experiences. *The Journal of Experimental Education*, 75(1), 5–29.
- Goetz, T., Nett, U. E., Martiny, S. E., Hall, N. C., Pekrun, R., Dettmers, S., & Trautwein, U. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. *Learning and Individual Differences*, 22(2), 225–234.
- Götz, T., & Hall, N. C. (2014). Academic boredom. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 321–340). New York/ London: Routledge/Taylor & Francis.
- Gravetter, F. J., Wallnau, L. B., Forzano, L.-A. B., & Witnauer, J. E. (2020). *Essentials of statistics for the behavioral sciences (10th ed.)*. Boston, USA: Cengage Learning.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis (7th ed.)*. New York: Pearson Education.
- Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes*, 38, 1–13.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132.
- Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326(5958), 1410–1412.
- Khajavy, G. H., MacIntyre, P. D., & Barabadi, E. (2018). Role of the emotions and classroom environment in willingness to communicate: Applying doubly latent multilevel analysis in second language acquisition research. *Studies in Second Language Acquisition*, 40(3), 605–624.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.)*. New York/London: Guilford Press.
- Kruk, M. (2016). Variations in motivation, anxiety and boredom in learning English in Second Life. *The EuroCALL Review*, 23(1), 25–39.
- Kruk, M. (2019). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in Second Life: the case of two advanced learners of English. *Computer Assisted Language Learning*, DOI: 10.1080/09588221.2019.1677722.
- Kruk, M., & Zawodniak, J. (2017). Nuda a praktyczna nauka je Rzyka angielskiego [Boredom and practical learning of English]. *Neofilolog*, 49(1), 115–131.
- Li, C. (2018). A positive psychology perspective on Chinese students' emotional intelligence, classroom emotions, and EFL learning achievement. Unpublished PhD dissertation, Xiamen University, China.
- Li, C. (2020). A Positive Psychology Perspective on Chinese EFL Students' Trait Emotional Intelligence, Foreign Language Enjoyment and EFL Learning Achievement. *Journal of Multilingual and Multicultural Development*, 41(3), 246-263. <https://doi.org/https://doi.org/10.1080/01434632.2019.1614187>
- Li, C. (2020). A Control-value Theory Approach to Boredom in English Class Among University Students in China. *The Modern Language Journal*. 105(1), in production.
- Li, C., Dewaele, J.-M., & Jiang, G. (2019). The complex relationship between

- classroom emotions and EFL achievement in China. *Applied Linguistics Review*.
<https://doi.org/10.1515/applirev-2018-0043>
- Li, C., Jiang, G., & Dewaele, J.-M. (2018). Understanding Chinese high school students' Foreign Language Enjoyment: Validation of the Chinese version of the Foreign Language Enjoyment scale. *System*, 76, 183–196.
- Liu, M., & Wronski, L. (2018). Trap questions in online surveys: Results from three web survey experiments. *International Journal of Market Research*, 60(1), 32–49.
- MacIntyre, P. D., Dewaele, J.-M., Macmillan, N., & Li, C. (2019). The emotional underpinnings of Gardner's Attitudes and Motivation Test Battery. In P. D. MacIntyre & A. Al-Hoorie (Eds.), *Contemporary language motivation theory: 60 years since Gardner and Lambert (1959)*. (pp. 57–79). Bristol: Multilingual Matters.
- Macklem, G. L. (2015). *Boredom in the Classroom: Addressing Student Motivation, Self-Regulation, and Engagement in Learning*. Manchester: Springer.
- Pawlak, M., Kruk, M., & Zawodniak, J. (2020a). Individual trajectories of boredom in learning English as a foreign language at the university level: insights from three students' self-reported experience. *Innovation in Language Learning and Teaching*. <https://doi.org/10.1080/17501229.2020.1767108>
- Pawlak, M., Kruk, M., & Zawodniak, J. (2020b). Investigating individual trajectories in experiencing boredom in the language classroom: The case of 11 Polish students of English. *Language Teaching Research*, 1–20.
<https://doi.org/10.1177/1362168820914004>
- Pawlak, M., Kruk, M., Zawodniak, J., & Pasikowski, S. (2020). Investigating factors responsible for boredom in English classes: The case of advanced learners. *System*, 102259.
- Pawlak, M., Zawodniak, J., & Kruk, M. (2020). The neglected emotion of boredom in teaching English to advanced learners. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12302>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341.
- Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass*, 4(4), 238–255.
- Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531–549.
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36–48.
- Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass*, 4(4), 238–255.
- Phakiti, A., De Costa, P., Plonsky, L., Starfield, S. (2018). Applied linguistics research: Current issues, methods and trends. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 5–29). New York, NY: Palgrave.
- Prior, M. T. (2019). Elephants in the room: An “affective turn,” or just feeling our way? *The Modern Language Journal*, 103(2), 516–527.
- Putwain, D. W., Pekrun, R., Nicholson, L. J., Symes, W., Becker, S., & Marsh, H. W.

- (2018). Control-value appraisals, enjoyment, and boredom in mathematics: A longitudinal latent interaction analysis. *American Educational Research Journal*, 55(6), 1339–1368.
- Ragheb, M. G., & Merydith, S. P. (2001). Development and validation of a multidimensional scale measuring free time boredom. *Leisure Studies*, 20(1), 41–59.
- Teimouri, Y. (2018). Differential Roles of Shame and Guilt in L2 Learning: How Bad Is Bad? *Modern Language Journal*, 102(4), 632–652.
- Vodanovich, S. J., & Watt, J. D. (2016). Self-report measures of boredom: An updated review of the literature. *The Journal of Psychology*, 150(2), 196–228.
- Zawodniak, J., & Kruk, M. (2019). Boredom in the English language classroom: An investigation of three language learners. *Konin Language Studies*, 7(2), 197–214.
- Zawodniak, J., Kruk, M., & Chumas, J. (2017). Towards conceptualizing boredom as an emotion in the EFL academic context. *Konin Language Studies*, 5(4), 425–441.
- Zuckerman, M. (1979). *Sensation seeking: beyond the optimal level of arousal*. Hillsdale, NJ Lawrence Erlbaum.

Appendix

Appendix 1 The validated *Foreign Language Boredom Scale* (Chinese and English versions)

The *Foreign Language Boredom Scale* (FLBS) is a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

No.	因子（分量表）和题项	Factors (Subscales) and Items	Factor loading
	因子 1 外语课堂无聊	Factor 1 <i>Foreign Language Class Boredom</i>	
1	英语课没什么意思。	The English class bores me.	.746
2	上英语课我容易犯困。	I start yawning in English class because I'm so bored.	.758
3	上英语课我容易发呆。	My mind begins to wander in the English class.	.860
4	英语课上我人在教室，思想却在神游。	I am only physically in the classroom, while my mind is wandering outside the English class.	.880
5	上英语课我很难集中注意力。	It is difficult for me to concentrate in the English class.	.860
6	英语课上时间过得好慢。	Time is dragging on in English class.	.747
7	英语课上我经常有听不下去的感觉。	I get restless and can't wait for the English class to end.	.834
8	英语课上我总是想办法打发时间。	I always think about what else I might be doing to kill the time rather than sitting in this English class.	.822
	因子 2 任务过简型无聊	Factor 2 <i>Under-challenging Task Boredom</i>	

9	我觉得长（英语）课文分析真的太枯燥了。	I believe an analysis of long text in English is really dreary.	.718
10	跟读（英语）课文好枯燥乏味。	It is really boring to repeat the (English) text after the modeling audio.	.749
11	太多类似的（英语）练习好没意思。	So many similar types of (English) exercises make me lose interest.	.838
12	对同一话题，太多重复的（英语）练习会让我很烦。	So much practice on a same (English-related) subject matter makes me restless.	.885
13	同一（英语）练习或内容持续时间太久，我觉得没意思。	The (English-related) exercise or a subject matter lasts too long, and I feel bored.	.857
	因子 3 PPT 情境型无聊	Factor 3 PowerPoint Presentation Boredom	
14	相比全是文字的 PPT，我觉得有其他多媒体资源会有意思些。	It would have been more interested if other multimedia resources were utilized in class rather than PPT slides loaded with text.	.822
15	全是文字没有互动的 PPT 让我很乏味。	PPT slides filled up with solely script but without interactions make me bored.	.882
16	“读 PPT”让我觉得很无聊。	Reading from script in the PPT slides bores me.	.745
	因子 4 作业情境型无聊	Factor 4 Homework Boredom	
17	想到英语作业就很烦躁。	Just thinking of my English homework makes me feel bored.	.833
18	英语作业太多了，我好烦。	I get bored of too much English homework.	.818
19	课后作业太难了，我不想做。	English homework is over-challenging and I don't want to do it.	.841
20	写作业是好枯燥的事情。	Doing homework is a dull activity.	.823
	因子 5 厌师型无聊	Factor 5 Teacher-dislike boredom	
21	英语老师不是我喜欢的那种类型（语音、语调或相貌），所以我对英语课也不是很感兴趣。	I am not interested in English class, because the English teacher isn't likable (e.g., tone, pitch or facial appearance).	.814
22	英语老师比较无趣，所以课堂比较沉闷。	The English teacher is an uninteresting, so the English class is dull.	.821
23	我真的不想听英语老师花很长的时间发表自己的个人感慨。	I really dislike the English teacher spending so much time making personal comments.	.696
24	英语老师花很长时间讲与教学内容无关的事情，我觉得好厌烦。	I feel agitated because the English teacher spends too much time saying things that are irrelevant to the teaching material.	.663

	因子 6 学习特质型无聊	Factor 6: <i>General Learning Trait Boredom</i>	
25	学习的时候,我总是感到无聊。	I'm always bored when I study.	.767
26	我就是对学习不太感兴趣的人。	I'm somebody who is not interested in study.	.830
27	学习(不仅仅英语学习)是一件无聊的事情。	Not only learning English, studying is dull in general.	.887
28	其他科目和英语一样无聊。	Other subjects are similarly boring and dull like English.	.781
29	我学习所有科目只是因为我不得不学。	I'm forced to learn all the subjects including English.	.772
	因子 7 任务过难或缺乏意义型无聊	Factor 7 <i>Over-challenging or Meaningless Task Boredom</i>	
30	英语老师不是很重视的环节,我也会觉得无所谓。	I don't care about teaching and learning activities that the English teacher does not value.	.747
31	英语老师不想讲的时候,我也不想听。	When the English teacher seems unmotivated to teach, I lose my motivation to listen to him/her as well.	.803
32	听不懂同伴的口语展示时,我会觉得很无聊。	If I cannot understand classmates' presentations, I become really bored.	.653