How teacher behaviour shapes Foreign Language learners' Enjoyment, Anxiety and Attitudes/Motivation: A mixed modelling longitudinal investigation[1]

Jean-Marc Dewaele*, Birkbeck, University of London
Kazuya Saito, University College London
Florentina Halimi, The Gulf University for Science & Technology

Corresponding Author*: Jean-Marc Dewaele, Department of Languages, Cultures and Applied Linguistics, 26 Russell Square, WC1B 5DT London, UK

Abstract
The current study investigates how Foreign Language Enjoyment (FLE), Foreign Language Classroom Anxiety (FLCA) and Attitude/motivation (AM) of 360 learners of English, German, French and Spanish in a Kuwaiti university was shaped over the course of one semester by three teacher behaviours: frequency of using the FL in class, predictability and frequency of joking. Linear mixed modelling revealed a positive relationship between the three teacher behaviours and FLE as well as AM, but no significant relationship emerged with FLCA. Multiple comparison analyses showed that levels of FLE dropped significantly among students whose teacher joked very infrequently and infrequently. It thus seems that the absence of teacher jokes had a delayed cumulative effect on FLE. No interaction effects were found with time for FLCA and for AM. We conclude that teacher behaviours affect both AM and FLE, and that teachers' absence of joking actually drains FLE over time.

Keywords: Foreign Language Enjoyment, Foreign Language Classroom Anxiety, Attitude/motivations, teacher Foreign Language use, teacher predictability, teacher joking

I Introduction

Interest in Foreign Language (FL) learner emotions has soared in the past decade (Dewaele, Chen et al. 2019). The fact that positive and negative emotions that learners experience in the classroom have a significant effect on their performance in the FL, as shown in the meta-analyses of Botes, Dewaele and Greiff (to appear), Botes, Greiff and Dewaele (2020) and Teimouri, Goetze and Plonsky (2019) means that they cannot be dismissed as mere froth in a field still dominated by cognitive perspectives on FL learning.

While learner emotions, and especially anxiety, were not completely absent in attitude and motivation research initiated by Gardner (1985)[1] and developed further by Dörnyei (2009), they appeared at item-level rather than at dimension-level (with the exception of anxiety). In other words, positive emotions were constituents of attitudinal and motivational dimensions but were not really considered in their own right. Although motivation researchers have increasingly acknowledged that emotions matter in FL learning, they still consider it to be background

---

variables that may undergird motivation but have little direct effect on FL learning. In his recent book on the future of language learning motivation, Dörnyei (2020) does not deny that learner emotions play an important role in FL learning but he does not quite agree with the view that emotions are motivating in themselves, arguing that they merely contribute to L2 motivational processes (p. 122).

One of the striking findings that have emerged from research on learner emotions is that FLE is mainly predicted by learner-external sources, such as teacher and peers, while FLCA is more linked to learner-internal sources such as personality traits (Dewaele & MacIntyre, 2019). Teachers who have their finger on the emotional pulse of the class can resort to a range of strategies to lift the learners' mood and stave off boredom. Some strategies to boost enjoyment and excitement are obvious and are often linked to teacher characteristics and their classroom management: introducing activities that are neither over- nor under-challenging, revealing the richness of the target culture, displaying enthusiasm about the language, the culture and the students in the classroom, being friendly, joking to promote group solidarity, encouraging and praising students, being well-organised and knowledgeable, establishing a relationship with students based on trust, mutual respect and common purpose (Dewaele & MacIntyre, 2014; 2019; Li, 2021). Neff and Rucynski (2017, 2021) showed that teachers who frequently use spontaneous jokes in the FL classroom manage to build a stronger rapport with their students and boost their motivation. Other strategies to create a more enjoyable class are less obvious, such as frequent use of the target language in class, and being more or less unpredictable – something that seems to be more appreciated in the Western than in the Asian context (Dewaele, Witney, Saito & Dewaele, 2018; Jiang & Dewaele, 2019).

Recent research has also started to include motivation constructs in the research design in order to establish the connection between learners' emotions and motivation (Dewaele & Proietti Ergün, 2020a, b; Saito et al., 2018; Pavelescu, 2019; Teimouri, 2017; Zhang & Tsung, 2021).

To date, the bulk of studies on learner emotions have used cross-sectional designs which can at best provide snapshots of dynamic processes. As Dörnyei (2007) pointed out, longitudinal research designs are the gold standard in applied linguistics, as they allow researchers to describe patterns of change and they explain causality better than correlation designs based on cross-sectional data sets.

The present study answers the call in Dewaele and Li (2020) for more longitudinal studies on learner emotions and for the use of more sophisticated statistical techniques to get a better grasp of the dynamic character of complex interactions between independent and dependent variables. This call has been answered with innovative studies published in the last two years by Elahi Shirvan and his team of Iranian researchers. More work is needed, more specifically, research on the effect of specific teacher behaviours (FL use in class, predictability and joking) on learners' FLE, FLCA and Attitude/Motivation at different points during one semester.

II Literature review

Considering the wealth of research on the role of various emotions and of motivation on FL learning, it is impossible to do justice to the field. We will thus focus on studies that linked FLE, FLCA and AM and that included teacher-centered variables in the research design.

1 The origin and the measurement of FLE

The paper that initiated research into FLE as a variable in its own right was the mixed methods study by Dewaele and MacIntyre (2014). The authors defined the concept of FLE as:

> "a complex emotion, capturing interacting dimensions of challenge and perceived ability that reflect the human drive for success in the face of difficult tasks, pleasure is considered simply an agreeable feeling. On the one hand, enjoyment occurs when people not only meet their needs, but exceed them to accomplish something new or even unexpected; on the other hand, pleasure is a simpler feeling that something likable is happening" (Dewaele & MacIntyre, 2016, pp. 216-217).

Dewaele and MacIntyre (2014) combined their FLE scale, consisting of 21 items reflecting positive emotions towards the learning experience, peers and teacher, with an 8-item short scale of Foreign Language Classroom Anxiety (FLCA) in an online questionnaire (see below). They collected quantitative and qualitative data from 1746 FL learners aged between 11 and 75 from all over the world. The finding of a moderate negative correlation between both emotions was interpreted as evidence suggest that they are not in a seesaw relationship, but operate as separate but related dimensions. Statistical analyses revealed that FLE and FLCA were linked to a number of sociobiographical variables. Older learners, female learners, learners with a high level of multilingualism, advanced mastery in the FL, feeling above average in their FL class, and university students rather than secondary school pupils reported more FLE and less FLCA (Dewaele and MacIntyre, 2014). Further analysis of the database used by Dewaele and MacIntyre (2014) was carried out by Botes, Dewaele and Greiff (2021) with the aim of developing a psychometrically sound 9-item Short Foreign Language Enjoyment Scale (S-FLES). Principal Component Analysis on the original FLE items revealed a higher order FLE factor with three dimensions: FLE Teacher, FLE Personal and FLE Social.

Higher levels of FLE have been linked to higher proficiency scores and better test results in the FL class (Li, 2020). A recent meta-analysis by Botes, Dewaele and Greiff (to appear) of 53 studies showed that FLE was moderately positively correlated with academic achievement in the FL and with self-perceived achievement in the FL. A stronger positive correlation was found between FLE and Willingness to Communicate.

2 The origin and measurement of Foreign Language Classroom Anxiety

Horwitz et al. (1986) developed the construct of Foreign Language Classroom Anxiety which reflected an individual's tendency to be anxious in the specific situation of language learning. Horwitz (2017) explained that "specific anxieties have characteristics of both trait and state anxieties. When individuals experience Language Anxiety, they have the trait of feeling state anxiety when participating in language learning and/or use. It is also likely that individuals who experience Language Anxiety would feel anxious simply thinking about language learning and/or use" (p. 33). Looking back at her original study, Horwitz concluded that "people who are generally anxious in their lives may be slightly more likely to be anxious in language learning. This finding also means that some anxious language learners do not experience a general tendency to anxiety in their daily lives" (Horwitz, 2017, p. 34).

Recent meta-analyses have shown that high levels of anxiety are negatively linked with FL progress and performance (Teimouri et al., 2019). Botes, Greiff and Dewaele (2020) focused specifically on 67 studies that used FLCA and found a moderate negative correlation between FLCA and reading, writing, listening, and speaking in the FL.

3 Origin and measurement of Attitudes/Motivation

Gardner's (1985) highly influential socio-educational model shaped the field's understanding of motivation in Second Language Acquisition. Newer models have been introduced such as the L2 Self (Dörnyei, 2009), but as Dörnyei (2019, p. xxi) points out, Gardner's model is "still relevant". Gardner (1985) investigated learners' attitudes and motivation to learn a foreign language taking into account the influence of the classroom, the school and the larger societal context. The socio-educational model consists of four broad dimensions, namely integrativeness, attitudes towards the learning situation, language anxiety and language attitudes and motivation, as well as 12 sub-dimensions. Gardner (2019) explained that his goal was "to clarify the underlying process linking affective variables to language achievement" (p. 6). He developed the Attitudes/Motivation Test Battery (AMTB)[2], consisting of 104 items, to gather information on "the student's affective reaction to the classroom environment, the cultural influences on the learner's reaction to acquiring attributes of the cultural community, anxiety reactions when called upon to use the language, and the effort, persistence and satisfaction associated with the process" (p. 11). The sub-dimensions include interest in FL, parental encouragement, motivational intensity, FL class anxiety, evaluation of teacher, attitude toward learning the FL, desire to learn the FL, attitudes toward speakers of the FL, integrative orientation, instrumental orientation, course evaluation and FL use anxiety. A shorter 12-item version exists, tapping into the various dimensions with a single item.

A meta-analysis of 75 studies by Masgoret and Gardner (2003) showed a strong positive relationship between motivation and FL achievement. The authors concluded that the motivated learner is somebody who "expends effort, is persistent and attentive to the task at hand, has goals desires, and aspirations, enjoys the activity, experiences reinforcement from success and disappointment from failure, makes attributions concerning success and/or failure, is aroused, and makes use of strategies to aid in achieving goals. That is, the motivated individual exhibits many behaviors, feelings, cognitions, etc., that the individual who is unmotivated does not" (Masgoret & Gardner, 2003, p. 128).

4 Sources of FLE and FLCA

Research into the sources of FLE and FLCA in addition to the initial ones uncovered in Dewaele and MacIntyre (2014) revealed quite similar patterns in different geographical contexts. Levels of FLCA were most strongly predicted by factors linked to learner-internal factors while FLE was mostly predicted by learner-external factors.

The first study to consider both learner-internal factors and learner-external factors was Dewaele et al. (2018). The authors collected data from 189 secondary school pupils in two elite London secondary schools who were mostly studying French, German and Spanish as FLs. Relatively few learner-internal variables predicted FLCA, with more advanced FL learners, who felt they were above group average in the FL class and who had more positive attitudes toward the FL reporting lower levels of FLCA. A larger number of independent variables predicted FLE. They included positive attitudes toward the FL and the FL teacher, frequency of use of the FL by the teacher, teacher unpredictability in class, i.e. the avoidance of set routines, the opportunity to speak up in class, and being more advanced in the FL. A follow-up study by Dewaele and Dewaele (2020) on a subsample of 40 students who had two teachers for the same FL revealed that FLCA was identical with the main teacher and the second teacher, but FLE was significantly lower with the latter. The positive predictors of FLE such as attitude toward the teacher, teacher's frequency of use of the FL in class, and unpredictability were significantly stronger for

the main teacher than for the second teacher. It strongly suggested that the teacher had the power to boost FLE but could do little to reduce FLCA.

A similar pattern emerged among 564 Chinese undergraduate EFL learners (Jiang and Dewaele, 2019). Participants' FLE was mainly predicted by teacher-related variables (attitudes towards the teacher, teacher's joking and friendliness – but not teachers' un/predictability- while FLCA was mainly predicted by learner-internal variables. Qualitative analysis of reports on episodes of high FLE and FLCA showed that the teacher was mentioned more frequently in relation to FLE than in relation with FLCA. Learners reported enjoying peer interaction, and smooth performance in answering teacher questions. Tests provoked spikes in FLCA. The crucial role of teachers in affecting learners' FLE was also found in Jiang (2020) who used the focused essay technique with 646 Chinese EFL students. FLE was positively linked with teacher friendliness, patience, kindness, happiness and regular use of humor.

Broadly similar relationships were found among 592 Kazakh learners of Turkish (Dewaele, Özdemir, et al., 2019). The strongest predictor of FLE was learners' attitude toward the FL followed by teacher's friendliness, teacher's higher frequency of use of the FL, teacher's lower strictness, and a more positive attitude toward the teacher (teachers' un/predictability was not included in the research design). In contrast with previous studies, FLCA was only weakly predicted by learner-internal variables: the strongest predictor was FL exam result, followed by attitude toward the FL, teacher friendliness and strictness (p. 15).

Using data from an international sample of 750 FL learners, Dewaele and MacIntyre (2019) found that attitude toward the teacher was the strongest predictor of FLE, followed by the personality trait Cultural Empathy, friendliness of the teacher, frequency of joking by the teacher, attitudes toward the FL, FL test results and low Social Initiative. Teacher predictability had no effect. In contrast, two personality traits, Emotional Stability and Social Initiative, were strong negative predictors of FLCA, followed by relative standing in the group, multilingualism, FL level, and attitude toward the teacher. A qualitative analysis of participants' narratives of classroom episodes in which they had experienced intense FLE and FLCA showed significant differences in frequency of mentions of teacher and the learner self. Teachers were mentioned in close to half the episodes related to FLE and only a quarter of the FLCA episodes. The learner self was mentioned close to 40% of the episodes of FLCA compared to less than a quarter of episodes of FLE. The quantitative and qualitative analyses thus independently confirmed that FLE and FLCA are conceptually different dimensions and are therefore not the two faces of Janus (Dewaele & MacIntyre, 2014).

Focusing exclusively on the effect of teacher-centred variables on FLE and FLCA, Dewaele, Franco Magdalena and Saito (2019) collected data from 210 EFL learners from Spain. Teacher friendliness was a strong predictor of FLE while a teacher's strong foreign accent in English was a negative predictor of FLE. Participants with younger teachers, who were very strict and did not use much English in class reported higher levels of FLCA.

Broadening the scope of the research, Zhang and Tsung (2021) focused on FLE of 216 adult international students studying Chinese in China. They found that FLE was mostly linked to a sense of personal fulfilment and a desire to progress in order to communicate in daily life. Interviews revealed that although their FLE was influenced by peers and teachers, the effect might have been more indirect than in Jiang and Dewaele (2019).

Li, Huang and Li (2020) decided to look not just at the effect of the teacher but at the role of the whole classroom environment on FLE and FLCA. Participants were 1718 secondary school students and 1295 university students studying English in China. They found that FLE

was more strongly predicted by classroom environment than by Trait Emotional Intelligence (TEI). The opposite pattern emerged for FLCA, which was more strongly predicted by TEI than by classroom environment. FLE was positively linked with teachers providing both academic and emotional support, who were passionate, and established a warm and caring atmosphere in the classroom.

Pursuing this path, Li et al. (2021) focused on the direct and indirect relationships between classroom environment, FLE, FLCA, FL boredom and Willingness to communicate (WTC) of 2268 EFL students in China. The three emotions were found to mediate the relationship between classroom environment and WTC in parallel, with FLE having the largest mediating effect, followed by FLCA.

The role of learner emotions and perceptions of teacher enthusiasm was at the heart of Dewaele and Li's (2021) study of 2002 Chinese EFL learners. The authors found that FLE and FL boredom co-mediated the relationship between teacher enthusiasm and student engagement. Interviews showed that teachers' verbal and/or nonverbal behaviors in the classroom shaped students' perceptions of their teaching enthusiasm.

Finally, Talebzadeh et al. (2020) focused on emotion contagion among five Iranian EFL learners using an idiodynamic research design. The authors found that enjoyment in face-to-face interaction was transmitted by teacher and student through mimicry of facial expressions, gesture, posture, and vocalization resulting in moment-by-moment variation in enjoyment.

5 Motivation and emotion

MacIntyre and Vincze (2017) investigated the relationship between 10 positive emotions (joy, gratitude, serenity, interest, hope, pride, amusement, inspiration, awe, and love) and 9 negative emotions (anger, contempt, disgust, embarrassment, guilt, hate, sadness, feeling scared, and being stressed) with motivation variables among 183 foreign language learners of German from Italian secondary schools in South Tyrol. The authors found that positive emotions were strongly correlated with motivation variables. The relationships between negative emotions and motivation were weaker and less consistent.

Pursuing a similar avenue, Zhang, Dai and Wang (2020) found that their 589 Chinese EFL students' motivation and language proficiency was mediated by FLE.

Similarly, Dewaele and Proietti Ergün (2020a, b) investigated relationships and differences in levels of classroom Enjoyment, Attitudes/Motivation (measured with Gardner's mini-AMBT) and Language Anxiety in 110 Turkish pupils' first language (L1), as well as their Italian and English (FLs) in an immersion school in Istanbul. Attitudes/Motivation were found to be significantly positively correlated with enjoyment in Turkish L1, Italian L2 and English L3, and negatively correlated with anxiety in Italian L2 and English L3.

6 Longitudinal research

Longitudinal research on FLE and FLCA is expanding rapidly. An early attempt is Dewaele and Dewaele (2017) who adopted a pseudo-longitudinal design to look at variation in FLE and FLCA between the start and the end of secondary education. Participants were the same 189 British learners that were used in Dewaele et al. (2018) but they were divided in three age groups: 12-13 year olds, 14-15 year olds and 16-18 year olds. FLCA was found to be similar in the three groups, suggesting long-term stability. FLE was also relatively similar in the three groups, with a slight drop for the 14-15 year olds and a slightly higher mean for the oldest group. Closer analysis revealed much more striking differences in the predictors of FLE and FLCA in the three

age groups. FLE was predicted by social standing in the youngest group and language level predicted their FLCA. Attitude towards the FL was the strongest predictor of FLE in the middle group while social standing in the group and language level were strong predictors of FLCA. The predictors changed again in the oldest group, where attitude towards the teacher was the only predictor of FLE, while social standing predicted FLCA (Dewaele & Dewaele, 2017, p. 18). It suggested that while levels of FLE and FLCA may seem quite stable over time, the predictors of FLE and FLCA do change over time.

Saito, Dewaele, Abe and In'nami (2018) combined a cross-sectional and a longitudinal analysis (over one academic term) of 108 Japanese high school EFL students to investigate the effect of motivation, FLE, FLCA, and learning experience on progress in comprehensibility in English. FLCA and Ideal L2 self were found to predict comprehensibility at the start of the study (by which point they had already had several years of EFL instruction). Students' L2 oral proficiency development during the term (measured via perceived comprehensibility) were found to be positively predicted by FLE, and more weakly predicted by Ideal L2 self, and negatively predicted by FLCA.

Also combining motivation and emotion variables in a single research design, Pan and Zhang (2021) carried out a longitudinal study into changes of FLE and FLCA among 55 Chinese EFL learners over a period of 14 weeks, linking them to motivation and personality traits. They found a drop in FLCA over the first four classes after which it stabilized, FLE went up slightly in the same period. Both emotions changed little over time but overall FLCA was more stable than FLE. FLE was found to be significantly positively correlated with motivated behavior or intended efforts (Criterion Measure), "Ideal L2 Self, Ought-to L2 Self, Family Influences, English Learning Attitudes, Cultural Interest, and Linguistic Self-confidence" (p. 12). In contrast, FLCA was only significantly negatively correlated with Criterion Measure and Ought-to L2 Self. Extraversion was positively linked with FLE and negatively with FLCA. Participants reported higher levels of FLE when teachers were friendly and encouraging, which contributed to a positive class atmosphere, which was further linked to "how peers interacted with each other and how close peers formed groups" (p. 18). Finally, motivational factors were found to be more closely related to the amount of change over time in FLCA as compared to FLE.

Li and Xu (2019) ran a six-week Positive Psychology intervention (training in Emotional Intelligence and diary reflection) with 56 Chinese vulnerable high school EFL learners in two classes. The design included an experimental and a control group. Training consisted in raising awareness of emotion in self and others, recognising emotions, generating positive emotions, understanding causes and consequences of emotions, expressing and regulating emotions appropriately. Comparing pre- and post-intervention data, the authors found that learners' FLE had increased significantly while their FLCA had dropped significantly.

A group of researchers around Elahi Shirvan has been particularly active in investigating change in FLE and FLCA over time among Iranian EFL learners. Elahi Shirvan and Taherian (2021) looked at change in FLE and FLCA among 367 university students during one semester using latent growth curve modelling. The authors found a significant increase in FLE and a significant decrease in FLCA at the end of the semester. However, there was a relatively large amount of heterogeneity in the patterns of FLE and FLCA. Initial levels of FLE and FLCA did not predict change.

A follow-up study by Elahi Shirvan, Taherian and Yazdanmehr (2021) used longitudinal confirmatory factor analysis-curve of factors model to trace changes in FLE among 437 Iranian EFL learners at four points of time with a two-week interval. FLE was found to increase at the

group level but there was a lot of inter-individual variation. In contrast with previous studies, the authors found that participants with lower FLE at the start experienced a steeper increase over time, which the authors attribute to learners' motivation, changing attitude to EFL learning and the supportive role of the teacher. The authors conclude that given the dynamic and complex nature of learner emotions, the field needs more research using sophisticated statistical models in order to capture the interactive effects of learner and teacher variables.

Adopting a completely different methodological approach, Elahi Shirvan and Talebzadeh (2020) used retrodictive qualitative modelling to explore four emerging archetypes of FLE and FLCA among 15 Iranian EFL students (high FLE and high FLCA; high FLE and low FLCA; low FLE and high FLCA; low FLE and low FLCA). Interviews with one prototypical student from each archetype reflecting on their past classroom emotions allowed the researchers to shed light on "the trends and trajectories leading to a specific outcome or attractor state by tracking and investigating the dynamic occurrences backwards" (p. 23). The authors conclude that the archetypes of FLE and FLCA "might stem from attractors including the influence of the teacher, personal goals, a perfectionist image of oneself and dissatisfactory and unsuccessful experiences in the past" (p. 40).

In a parallel but totally distinct study based on the same database of 360 FL learners from Kuwait as used in the present study, Dewaele, Saito and Halimi (2022) used Repeated measures ANOVAs to measure change over time and mixed effects modelling to capture complex interactions between FLE and its three dimensions (FLE teacher, FLE Personal and FLE Social), FLCA, and Attitudes/Motivation and its four dimensions (Integrativeness, Attitudes toward the Learning Situation, Motivation and Instrumental orientation) over the course of one semester. The analyses revealed that FLE (including FLE Personal and FLE Social) and FLCA remained stable over time but that FLE Teacher decreased significantly. Similarly, levels of Attitudes/Motivation and more specifically the dimensions *Motivation* and *Attitudes toward the Learning Situation* decreased significantly as the term progressed. FLE (including FLE Personal and FLE Social) and FLCA were found to have medium effects on Attitudes/Motivation dimensions. Significant interaction effects for FLE Teacher and Time for the Attitudes/Motivation dimensions were interpreted as evidence that FLE can act as a buoy for sagging motivation and thus become intrinsically motivating. The study did not include teacher behaviour variables, which are the object of the present study.

6 Rationale of the present study

What the literature review has shown is that there are complex interactions between learner-internal and learner-external variables that affect levels of FLE and FLCA directly and indirectly. The majority of the research is based on cross-sectional designs, which allows the identification of sources of FLE and FLCA but cannot attribute causality. Indeed, emotions can shape behaviour but the behaviour (such as performing a task un/successfully) also influences the emotions, so there is a clear feedback loop. The recent surge in longitudinal research designs allow researchers to get closer to causality by zooming in on changes in FLE and FLCA over time but also on potential change in predictors of FLE and FLCA. The current study will pursue this path, using mixed modelling to analyse the evolving relationship between the dependent variables and three independent variables that have shown to be strong predictors of FLE: teachers' frequency of FL use in class, teacher predictability and frequency of joking in class by the teacher. More specifically, we explore the following two related research questions:

- To what extent do learners' FLE, FLCA, and AM (Attitude/Motivation) change over four data collection points during one semester of FL teaching?
- How is this change linked to teachers' behavior (frequency of FL use, predictability and frequency of joking)?

III Methodology

1 Participants

Participants were 360 university students (280 females, 75 males, 5 did not say) in Kuwait enrolled in FL classes, ranging from beginner to advanced, including English (n = 252), German (n = 43), Spanish (n = 41) and French (n = 24). All participants had Arabic as a first language, sometimes combined with other first languages. The sample consisted of 177 bilinguals, 125 trilinguals, 37 quadrilinguals, 14 pentalinguals and seven participants reported knowing six or more languages. A large majority of participants were Kuwaiti (n = 317, 88%) with increasingly smaller numbers of other nationalities, including Lebanese (n = 8), Jordanian, Iranian, Syrian, Egyptian, Saudi Arabian, Canadian, Indian, Palestinian, American, Bahraini, Iraqi, Pakistani and Swedish. Mean age was 20 years (SD = 3.25), with a range between 17 and 48, and 96% of the sample being under the age of 26.

2 Institution

The Gulf University for Science & Technology (GUST) offers a Western-style education "grounded in Kuwait's Islamic cultural setting" with a large majority of the student population being Kuwaiti (https://www.higheredjobs.com/InstitutionProfile.cfm?ProfileID=15942). GUST is linked with the University of St. Louis – Missouri and offers an American-style accredited degree program with English as the medium of instruction. Competence in English is thus a prerequisite for academic success. The mission statement of the GUST Foundation Unit, the English Department and the Department of Humanities and Social Sciences states that teachers adopt a "task-based approach to provide students with a command of modern languages and cultural competencies as well as critical thinking and communication skills that allow for interaction in real-life settings" (https://www.gust.edu.kw/content/languages_0) The curriculum includes attention to grammar, culture and communication.

3 Instruments

The questionnaire was presented in English, a language mastered to at least an intermediate level by all students. A first set of questions focused on participants' sociobiographical background and their language learning history, such as the languages known, age and gender. This was followed by items for FLE, FLCA, Attitudes/Motivation in the Foreign Language class. Foreign Language Enjoyment (FLE) was measured using Botes et al.'s (2021) 9-item short-form questionnaire. Items reflect three dimensions: Personal Enjoyment ('I am proud of my accomplishments'), Social Enjoyment ('We laugh a lot'), and Teacher Appreciation ('The teacher is encouraging'). Responses to these items were given on a 5-point Likert scale (1 = 'strongly disagree', 2 = 'disagree', 3 = 'undecided', 4 = 'agree', 5 = 'strongly agree'. All items were positively phrased. Scale analyses revealed high internal consistency (see Table 1).

Foreign Language Classroom Anxiety (FLCA) was measured with the eight-item scale extracted from Horwitz et al.'s (1986) original 33-item scale previously used in Dewaele and MacIntyre (2014) and validated in Botes et al. (to appear). Items were accompanied by a 5-point

Likert-type scale (from 1 = 'strongly disagree' to 5 = 'strongly agree'). They referred to mild and more severe physical symptoms of anxiety ('I can feel my heart pounding when I'm going to be called on in FL class') and to nervousness and self-confidence ('I get nervous and confused when I am speaking in my FL class'). The two positively phrased items were reverse-coded so that a higher score reflects a higher level of FLCA. Reliability statistics were very good (see Table 1).

Attitudes and motivation (AM) was measured with the slightly adapted 11-item short version of Gardner's (1985) Attitudes/Motivation Test Battery (AMTB) with 7-point Likert scales, with anchors including "weak to strong", "unfavourable" to "favourable", "very low/little" to "very much/high". In this short instrument, "each item corresponds to one of the 11 scales in the larger battery" (Tennant & Gardner, 2004, p. 247). It allows researchers to calculate scores for Integrativeness (integrative orientation + interest in Foreign Languages + attitude toward French Canadians), Attitudes toward the Learning Situation (attitude toward French instructor + attitude toward French course), Motivation (motivational intensity + desire to learn the FL + attitude toward learning French) and Language Anxiety (French use anxiety + French course anxiety), Instrumental orientation (instrumental orientation). Anxiety was left out because it would overlap with FLCA, which is an independent variable. Reliability statistics were very good (see Table 1). The adaptation involved turning the statements into generic ones, removing references to French and the Canadian context. We opted for Gardner's framework rather than the more recent L2 Self framework developed by Dörnyei (2009) because we are particularly interested in the effect of one specific contextual factor, teacher behaviour, not on students' L2 Ideal self nor their Ought-to-L2 selves, but on their attitudes and orientation towards the FL and its speakers, towards the language learning situation, and the strength of their desire to acquire the FL. Moreover, we wanted to include the wider institutional and cultural context in which our Arab participants learn their FLs. We reported earlier that Dörnyei (2019) declared that Gardner's model may have had a different focus than the L2 Self approach but that it was still valid.

A series of QQ plots showed that the dependent variables followed a normal distribution pretty well (available from the authors on request).

Table 1. Internal consistency of FLE, FLCA and AM at Time 1, Time 2, Time 3 and Time 4 (Cronbach's alpha)

Time	FLE	FLCA	AM
T1	.810	.766	.874
T2	.889	.784	.933
T3	.906	.801	.945
T4	.927	.832	.949

4 Data collection

Data were collected four times between October and December 2019 in the FL classes. The semester started in September, with preparations for the mid-term exams in October. To collect the data, the survey was created with the LimeSurvey web application along with calendaring software. The faculty teaching English, German, Spanish, and French were invited to a meeting where a complete overview of the study was presented. After that, the faculty and one of the researchers met the students, explained the study and its purpose, presented the consent form,

which guaranteed privacy, anonymity, and the right to participate or drop out, and extended the invitation to participate in the study. Consent was given online by ticking the corresponding box. The survey link was sent four times (at two-week intervals) to the students attending the English, German, Spanish and French courses. They were asked to submit the survey by a given date.

5 Analytical strategy

To examine the source of individual variation among students' FLE, FLCA and AM from multiple angles, the present study will use linear mixed models. Linear mixed models allow both fixed and random effects, and are particularly useful when there is non-independence in the data, i.e., both dependent and independent variables collected from the same group of participants over time. To unpack the longitudinal relationship between the dependent variables and the teacher behaviour variables, three different linear mixed-effects models were constructed using the lme4 package in the R environment (version 1.1-23, Bates et al., 2015). In each model, one of students' emotion/motivation variables (FLE, FLCA, and AM) serves as dependent variable (Models 1, 2, and 3, respectively). Both time (T1-T4) and the teacher behaviour variables were entered as fixed-effects predictors. Time was treatment-coded. Participants were entered as random effects. Featuring both main and interaction effects of Time, FL Use, Predictability, and Jokes, the model was formulated as follows:

- Dependent Variables (FLE, FLCA, or AM) = Time + FL Use + Predictability + Jokes + Time × FL Use + Time × Predictability + Time × Jokes

Following Westfall, Kenny and Judd's (2014) recommended procedure for calculating statistical power for a crossed design model, the sample size (360 participants × 4 data collection points = 1840 observations) generated .951 in order to obtain a medium effect size ($d = 0.5$). The figure substantially exceeds Larson-Hall's (2015) field-specific benchmark of $\alpha = .700$. The descriptive statistics for dependent and independent variables are available in the appendix. Where significant main and interaction effects were found, we conducted follow-up analyses.

6 Research ethics

Data was collected after getting approval from the University's Institutional Review Board (IRB)[3]. The University's IRB approval was also mentioned in the consent.

IV Results

1 Model 1: Foreign Language Enjoyment

As summarized in Table 2, the model yielded a range of significant main effects of teacher behaviour variables ($p < .001$ for FL Use, Predictability, and Jokes). Throughout the project, participants who demonstrated greater enjoyment likely had teachers who used FL more often, delivered more predictable lessons, and produced more jokes. Given that main effects of Time reached statistical significance ($p < .001$), the participants' FLE appeared to change to some degree. According to post-hoc multiple comparison analyses, surprisingly, the participants' FLE scores were significantly lower at T4 than T1 ($M = 4.01$ vs. 3.89, $t = 2.356$, $p = .019$, $d = 0.95$). In terms of the source of individual differences, it is worth pointing out that the interaction effect

of Time and Jokes reached statistical significance ($p < .001$); but that the interaction effects of Time and the other two teacher variables (FL Use, Predictability) failed to reach statistical significance, $p < .384, .838$). The results indicate that participants' FLE decline could be related to Jokes but not to FL Use and Predictability. To this end, we conducted post-hoc analyses on the significant interaction effects of Time and Jokes while no other follow-up analysis was done on the non-significant interaction effects of Time, FL Use, and Predictability.

Table 2. Summary of Mixed-Effects Regression Analysis of FLE Scores Relative to Time and Teacher Quantity Variables (Model 1)

	Fixed effects: Factor	*Estimate*	*SE*	*t*	*p*
FLE scores (T1-T4)	Intercepts	3.0200	0.2207	13.6830	< .001*
	Time	-0.2505	0.0745	-3.3640	< .001*
	FL Use	0.0779	0.0356	2.1880	< .028*
	Predictability	0.0767	0.0312	2.4560	< .014*
	Jokes	0.0146	0.0265	0.5530	.580
	Time × FL Use	0.0111	0.0127	0.8710	.384
	Time × Predictability	-0.0024	0.0118	-0.2040	.838
	Time × Jokes	0.0382	0.0102	3.7450	< .001*
	Random effects	Variance	*SD*		
	Participants	0.008	0.091		
	Residual	0.336	0.580		
	Conditional R^2	.280			
	Marginal R^2	.262			

* indicates $p < .05$

To further investigate the complex relationship between FLE decline and the amount of teacher jokes, participants' mean scores were calculated as per the amount of teacher jokes (*7 = very frequent, 6 = frequent, 5 = rather frequent, 4 = mid, 3 = rather infrequent, 2 = infrequent, 1 = very infrequent*) at the four different time points (T1-T4). As visualized in Figure 1, while those who often observed teacher jokes maintained FLE without much change, FLE appeared to decline among those whose teacher joked less frequently. According to a two-way repeated ANOVA (Time × Group), the results demonstrated not only significant main effects of Time (*F* (1, 355) = 6.182, *p* = .013) but also significant interaction effects of Time and Group (*F* (5, 353) = 2.424, *p* < .001). The results of multiple comparison analyses showed that in the groups with Very Infrequent and Infrequent joking, levels of FLE significantly declined from T1 (*M* = 3, 66, 3.93) to T4 (*M* = 2.77, 2.96), *p* < .001, *d* = 1.20, 1.32, whereas FLE changes among the other groups did not reach statistical significance (*p* > .05). Taken together, the results suggest (a) that FLE is relatively stable as long as teachers make their classrooms fun with ample FL use, predictable FL activities, and frequent jokes; and (b) it may decline especially if teachers do not make enough jokes.
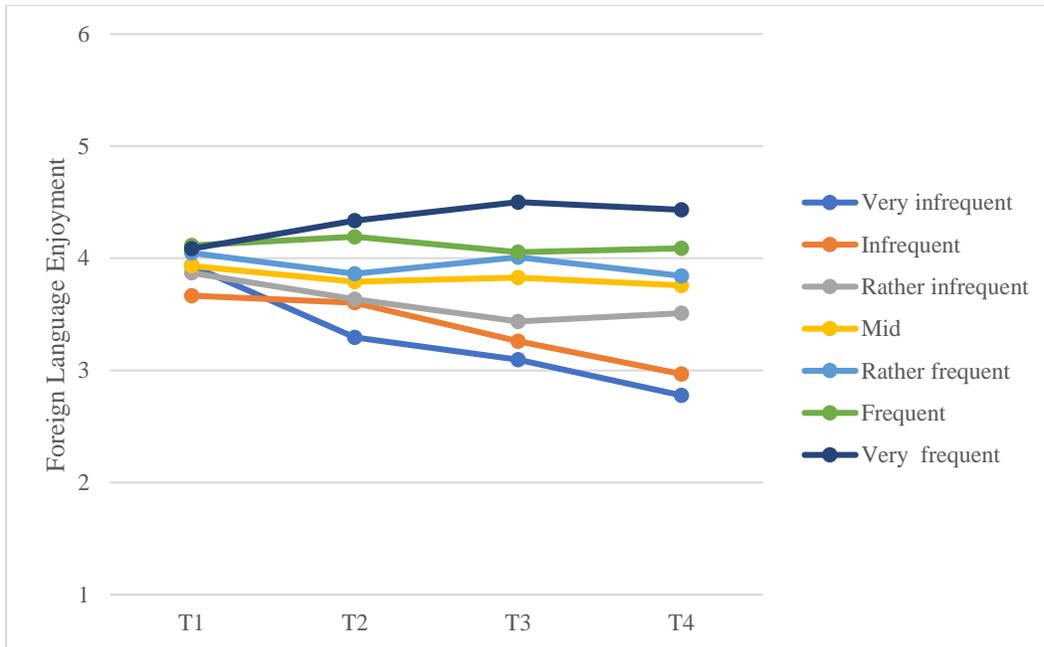
Figure 1. Amount of Teacher Joke Frequency (x axis) and Students' FLE (y axis) over Time

2 Model 2: Foreign Language Classroom Anxiety

Another set of mixed-effects model was constructed with FLCA as dependent variable relative to four predictor variables (Time, FL Use, Predictability, Jokes). As summarized in Table 3, The model did not find any significant main and effects of Time and the three teacher variables (FL Use, Predictability and Jokes ($p < .001$). The lack of significant main/interaction effects demonstrated that participants' FLCA scores were comparable across four different time points (M = 2.73, 2.75, 2.76, and 2.72). Also, participants' FLCA was unrelated to the extent to which teachers used a target language in classrooms, how much their class was predictable vs. unpredictable, and how much teachers used jokes.

Table 3. Summary of Mixed-Effects Regression Analysis of FLCA Scores Relative to Time and Teacher Quantity Variables (Model 2)

	Fixed effects: Factor	*Estimate*	*SE*	*t*	*p*
FLCA scores (T1-T4)	Intercepts	3.4390	.221	12.4430	< .001
	Time	-0.1668	.0745	-1.7880	.073
	FL Use	-0.0871	.0356	-1.9540	.051
	Predictability	-0.0689	.0312	-1.7650	.077
	Jokes	0.0501	.0265	1.5150	.130
	Time × FL Use	0.0064	.0127	0.4010	.688
	Time × Predictability	0.0161	.0118	1.0890	.276
	Time × Jokes	0.0051	.0102	0.4010	.688
	Random effects	Variance	*SD*		
	Participants	0.005	0.078		
	Residual	0.538	0.733		
	Conditional R^2	.030			
	Marginal R^2	.028			

* indicates $p < .05$

3 Model 3: Attitudes/Motivation

A final set of mixed-effects model was constructed with AM as dependent variable and four variables as predictors (Time, FL Use, Predictability, Jokes). Although the model yielded significant main effects of FL Use, Predictability, and Jokes ($p < .001$), those of Time were not significant ($p = .470$) (see Table 4). The results here (significant main effects but non-significant interaction effects) indicated that whereas participants' AM seemed to remain quite stable over time (different from FLE and similar to FLCA), their AM scores (at any time points) were significantly determined by teachers' efforts to use FL, make their classes predictable, and make jokes.

Table 4. Summary of Mixed-Effects Regression Analysis of AM Scores Relative to Time and Teacher Quantity Variables (Model 3)

	Fixed effects: Factor	*Estimate*	*SE*	*t*	*p*
AM scores (T1-T4)	Intercepts	1.606	0.3304	4.8610	< .001*
	Time	-0.1735	0.1115	-1.5570	.119
	FL Use	0.3694	0.0533	6.9260	< .001*
	Predictability	0.2126	0.0467	4.5480	< .001*
	Jokes	0.0873	0.0396	2.2030	.027*
	Time × FL Use	0.0081	0.0191	0.4260	.670
	Time × Predictability	0.0153	0.0177	0.8640	.387
	Time × Jokes	0.0158	0.0153	1.0300	.303
	Random effects	Variance	*SD*		
	Participants	0.02794	0.1671		
	Residual	0.74677	0.8642		
	Conditional R^2	.537			
	Marginal R^2	.519			

* indicates $p < .05$

V Discussion

The first research question focused on the amount of change in Arab FL learners' FLE, FLCA and AM levels over the course of one semester. Levels of FLCA and AM did not vary significantly over time while levels of FLE dropped towards the end of the semester. Dewaele, Saito and Halimi (2022) on the same dataset as in the present study showed that while dimensions *FLE Personal* and *FLE Social* remained stable, the values for the dimension *FLE Teacher* dropped significantly, explaining the decrease in overall FLE. The relative stability of learners' emotions over time broadly confirms patterns uncovered in some previous studies that did not involve some type of intervention or classroom observation (Dewaele & Dewaele, 2017; the control group in Li & Xu, 2019; Pan & Zhang, 2021). Change is typically linked to intervention or continued observation. For example, Li and Xu (2019) found increased FLE and lower FLCA in their experimental group at the end of the intervention during which learners were explicitly taught about emotion recognition and regulation. Weaker change has been observed when participants and their teachers are simply observed by the researchers. Elahi Shirvan et al. (2021), for example, reported an increase in FLE over time, which the authors attribute to learners' more positive attitudes and increased motivation resulting from visualisation of their ideal L2 selves during the FL course. The authors argue that teachers' emotional support may also play a role in the increase in FLE, though they did not measure whether this actually changed over time. They did observe that teachers were increasingly supportive during the classes (personal communication). It is possible that the continued presence of the researchers pushed both students and teachers to adopt their best behaviour, resulting in a positive emotional loop. This phenomenon is known as social desirability bias (Dörnyei, 2010). Elahi Shirvan and Taherian (2021) showed a significant decrease in FLCA and an increase in FLE without apparent cause. A closer look at the results show that the change had been significant but that the *p* value was just .05. This suggests that the effect size was probably very small. Analysis of the qualitative data of four participants in the same study, including journals, interviews and motometers that allowed participants to record their levels of FLE and FLCA every ten minutes, showed a highly dynamic picture, with spikes and dips in FLE and

FLCA linked to positive or negative teachers' feedback, peers' judgments and their own performance in tasks. This finding confirms the overall view that group averages of learner emotions hide a huge amount of individual variation (Elahi Shirvan & Talebzadeh, 2020; Elahi Shirvan et al., 2021; Pan & Zhang; 2021; Talebzadeh et al., 2020). The study by Dewaele et al. (2022) on the same dataset as in the present study showed that while levels of FLCA and FLE remained quite stable over time, there was a significant drop in motivation levels. The dimensions *Motivation* and *Attitudes toward the Learning Situation* dropped most significantly. FLCA was found to drag AM down while FLE supported AM, protecting it from a disappointing teacher or curriculum.

The second research question focused on the effect of different teacher behaviours on learners' FLE, FLCA and AM. The mixed-effects analyses on FLE revealed significant main effects of all three teacher behaviour variables ($R^2 = 26.2\%$). Throughout the semester students reported higher levels of FLE with teachers who used FL frequently, were predictable and joked frequently in class. This finding confirms earlier research on the central role of teachers in their learners' FLE (Dewaele et al., 2018, Dewaele and MacIntyre, 2019, Dewaele & Dewaele, 2020; Elahi Shirvan & Talebzadeh, 2020; Elahi Shirvan et al., 2021; Li et al., 2020; Jiang, 2020; Pan & Zhang; 2021; Talebzadeh et al., 2020). Further analysis of our data suggested that the drop in FLE was particularly pronounced among students who reported that their teacher did not joke much at all. This suggests that teacher behaviour can have a **delayed cumulative effect**. The lack of joking may not affect students' FLE at the start of the course but may start denting FLE after a while. Joking may in fact be only part of a wider picture. One could speculate that teachers who don't joke may also take themselves too seriously, are arrogant, unsmiling, unsecure, make a fuss about trivial issues, are too strict, too picky, not friendly enough, not generous in their marking of assignments and may fail to create a positive classroom atmosphere. Students' patience with this type of teacher may wear thin after a while and their FLE may drop as a consequence. It is probably no coincidence that it was the values on the dimension *FLE Teacher* that dropped significantly in Dewaele et al. (2022). On the other hand, teachers who use humour in the FL classroom, and especially spontaneous jokes, may in fact strengthen teacher-student rapport and create an environment in which learners feel motivated (Neff & Rucynski, 2017, 2021). The teachers in Neff and Rucynski (2021) reported that learner proficiency was not an obstacle to the use of humor and that they adjusted their humor so their learners would understand it. They insisted that humor was even more important for beginners than for more advanced students.

The mixed-effects analyses on FLCA revealed that teacher behaviours, namely FL use in classrooms, predictability and frequency of joking, had no significant effect on learners' FLCA. This confirms that teachers are limited in what they can do to lower students' FLCA in traditional classes (Dewaele & Dewaele, 2020). FLCA could be considered as a stable phenomenon that L2 leaners likely establish after years of classroom experience regardless of the immediate effects of their current experience with teachers and peers (Dewaele & Dewaele, 2017; Horwitz, 2017, Horwitz et al., 1986).

The final series of mixed-effects analyses revealed that AM remained unchanged over time and that it was strongly positively linked to teachers' frequency of FL use, joking and predictability ($R^2 = 51.9\%$). These behaviours could be seen as teacher strategies to generate initial motivation, and then to maintain and protect it (Dörnyei, 2001; Masgoret & Gardner, 2003). The similarity between the effects of the teacher behaviours on AM and on FLE is confirmed by previous research (Dewaele & Proietti Ergün, 2020a, b, 2020a, b; MacIntyre &

Vincze, 2017; Zhang & Tsung, 2021) and lends support to Dörnyei's (2020) argument that learner emotions can sustain and amplify their motivation. The finding that predictability was positively linked to AM may seem slightly surprising but the positive link with FLE diverges completely from some previous studies where predictability was linked to **lower** FLE (Dewaele & Dewaele, 2017, 2020; Dewaele et al., 2018) or where un/predictability was unrelated to FLE (Dewaele & MacIntyre, 2019; Jiang & Dewaele, 2019). Research outside applied linguistics has shown that appreciation of predictability depends on the individual, the situation and the larger social group. Individuals with autism spectrum disorder have an above-average preference for predictability (Goris et al., 2020). There are also situations where the preference for predictability predominates and where a person's unpredictable or inappropriate behavior will result in that person being liked less (Kiesler, 1973). However, unpredictability and norm violation can lead to a greater liking if that person is deemed interesting (Kiesler, 1973, p. 359). Dewaele et al. (2018) acknowledged that the participants did not represent the general population of FL learners as they were self-selected pupils from elite secondary schools, in other words, they were probably well above average in terms of motivation and skills, which may explain why they felt comfortable with less predictable teachers. Average, or below average, students in that population may have preferred more predictable teachers. We realise that more general sociocultural factors may also determine students' appreciation of unpredictable behavior. The current group of Arab learners probably differs from samples in previous studies in terms of religiosity. It is possible that through their daily prayers these Muslim learners have learned to appreciate routines, both in the mosque and in FL classrooms.

It is important to point out that while teachers have the power to affect the FLE, FLCA and AM of their students, there are many factors beyond their control (Talebzadeh et al., 2020). These can include learner-internal variables, like a shift in interest or a lack of time for homework, or learner-external variables such as group dynamics, where a learner may feel left out or not appreciated by peers, or where one student might be disruptive or unpleasant, spoiling the whole class atmosphere. Even experienced teachers struggle to neutralise the nefarious influence of such disruptors.

Causality is never entirely unidirectional in the context of emotion. Teachers can regulate classroom emotions through a process of emotional contagion (Dewaele & Li, 2021; Talebzadeh et al., 2020). They may feel bolstered by learners' positive reactions to their jokes and light-hearted remarks, and hence add a few, thus creating a positive feedback loop. Similarly, teachers may notice that their switching to another language in class, or their introduction of something unpredictable, or the absence of jokes leads to a cooling of the classroom atmosphere which might nudge them to return to their usual practice.

The current study is not without limitations. Relying exclusively on longitudinal quantitative data allowed us to establish changes over time and the growing effect of specific teacher behaviours but without participant voices or classroom observation, we have no independent way to establish causality. Although our mixed effects analyses indicate "effects", there is no way to confirm "effects" unless we conduct a proper intervention study to see whether increased joking in an experimental group would lead to higher levels of FLE than in a control group. It would be interesting then to observe the effect of different types of humour teachers use, whether this changes over time, and whether some particular types might be more appreciated than others, depending on their level of proficiency or the teacher being a L1 user of the language taught (cf. Neff & Rucynski, 2017, 2021). A second limitation is the fact we had only four data collection points, compared to the 43 data collection points in Pan and Zhang

(2021) for example. One consequence of this is a loss of granularity and an averaging out of values.

VI Conclusion

The present study set out to investigate how FL students' FLE, FLCA and AM was shaped by teachers' frequency of use of the FL (English, French, German and Spanish) in class, the teachers' predictability and the frequency of teachers' joking over the course of a semester. The results of mixed effects analyses showed a positive relationship between the three teacher behaviours and FLE as well as AM, and a negative relationship with FLCA. In other words, these three teacher behaviours play a crucial role in creating a positive emotional environment in which students can flourish. The most striking finding was a drop in FLE among students who had teachers who joked very infrequently. We argued that the lack of joking may have had little effect on students' FLE at the start of a course but may have had a delayed cumulative effect, just like the absence of yeast in the preparation of bread means the dough will not rise, which will affect the shape and the texture of the final bread. The metaphor of FL teachers being like bakers could be extended further. Just as every group of students is unique in its composition, so are the ingredients needed to make bread, including the size of the flour, the quality of the water, and the type of yeast. Depending on these differences, teachers and bakers may need to adjust their approach to create the optimal conditions for a good result.

ORCID iD
Jean-Marc Dewaele https://orcid.org/0000-0001-8480-0977
Kazuya Saito https://orcid.org/0000-0001-8065-0772
Florentina Halimi https://orcid.org/0000-0002-5966-830X

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Botes, E., Dewaele, J.-M., & Greiff, S. (2021). The Development of a Short-Form Foreign Language Enjoyment Scale. *The Modern Language Journal, 105*(4), 858-876.

Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: An overview of the literature and a preliminary meta-analysis of Foreign Language Enjoyment and other individual difference variables. Unpublished manuscript. DOI: 10.31234/osf.io/guaj5 Link: https://psyarxiv.com/guaj5

Botes, E., Greiff, S., & Dewaele, J.-M. (2020). The Foreign Language Classroom Anxiety Scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *The Journal for the Psychology of Language Learning, 2*, 26–56.

Botes, E., van der Westhuizen, L., Dewaele, J.-M., MacIntyre, P.D., & Greiff, S. (to appear). Validating the Short-Form Foreign Language Classroom Anxiety Scale (S-FLCAS). *Applied Linguistics*

Dewaele, J.-M., Chen, X., Padilla, A.M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in Psychology. Language Sciences, 10*, 2128.

Dewaele, J.-M., & Dewaele, L. (2017). The dynamic interactions in Foreign Language Classroom Anxiety and Foreign Language Enjoyment of pupils aged 12 to 18: A pseudo-longitudinal investigation. *Journal of the European Second Language Association*, *1*, 11-22.

Dewaele, J.-M., & Dewaele, L. (2020). Are foreign language learners' enjoyment and anxiety specific to the teacher? An investigation into the dynamics of learners' classroom emotions. *Studies in Second Language Learning and Teaching*, *10*, 45-65.

Dewaele, J.-M., Franco Magdalena, A., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL Learners' Anxiety and Enjoyment. *The Modern Language Journal*, 103(2), 412-427.

Dewaele, J.-M., & Li, C. (2020). Emotions in Second Language Acquisition: A critical review and research agenda. *Foreign Language World* [Chinese 外语界], *196*(1), 34-49.

Dewaele, J.-M. & Li, C. (2021). Teacher enthusiasm and students' social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes. *Language Teaching Research.* Epub ahead of print. DOI:10.1177/13621688211014538

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 Teaching*, *4*, 237-274.

Dewaele, J.-M., & MacIntyre, P.D. (2016). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of FL learning? In P.D. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215-236). Bristol: Multilingual Matters.

Dewaele, J.-M., & MacIntyre, P.D. (2019). The predictive power of multicultural personality traits, learner and teacher variables on Foreign Language Enjoyment and Anxiety. In M. Sato & S. Loewen (Eds.), *Evidence-based second language pedagogy: A collection of Instructed Second Language Acquisition studies* (pp. 263-286). London: Routledge.

Dewaele J.-M., Özdemir C., Karci D., Uysal S., Özdemir E. D., & Balta N. (2019). How distinctive is the Foreign Language Enjoyment and Foreign Language Classroom Anxiety of Kazakh learners of Turkish? *Applied Linguistics Review*. Epub ahead of print. DOI:10.1515/applirev-2019-0021.

Dewaele, J.-M., & Proietti Ergün, A. L. (2020a). Appagamento, atteggiamente/motivazione e ansia nello studio della lingua madre e della lingua stranieri in una scuola italiana all'estero. *Rassegna Italiana di Linguistica Applicata*, *51* (2-3), 203-220.

Dewaele, J.-M., & Proietti Ergün, A. L. (2020b). How different are the relations between enjoyment, anxiety, attitudes/motivation and course marks in pupils' Italian and English as Foreign languages? *Journal of the European Second Language Association*, *4*(1), 45–57.

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety in the FL classroom: The effect of teacher and learner variables. *Language Teaching Research*, *22*(6), 676-697.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom.* Cambridge: Cambridge University Press.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative and Mixed Methodologies.* Oxford: Oxford University Press.

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 9-42). Bristol: Multilingual Matters.

Dörnyei, Z. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing.* New York: Routledge.

Dörnyei, Z. (2019). Foreword. In A. Al-Hoorie & P. D. MacIntyre (Eds.), *Contemporary language motivation theory: 60 years since Gardner and Lambert (1959)* (pp. xix-xxii). Bristol: Multilingual Matters.

Dörnyei, Z. (2020). *Innovations and Challenges in Language Learning Motivation*. New York: Routledge.

Dörnyei, Z., & Ryan, S. (2015). *The Psychology of the Language Learner Revisited.* New York: Routledge.

Elahi Shirvan, M., & Taherian, T. (2021). 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*, *24*(1), 31-49.

Elahi Shirvan, M., & Talebzadeh, N. (2020). Tracing the signature dynamics of foreign language classroom anxiety and foreign language enjoyment: a retrodictive qualitative modelling. *European Journal of Applied Linguistics*, *6*, 23–44.

Elahi Shirvan, M., Taherian T., & Yazdanmehr, E. (2020). The Dynamics of Foreign Language Enjoyment: An Ecological Momentary Assessment. *Frontiers in Psychology*, *11*, DOI: 10.3389/fpsyg.2020.01391

Elahi Shirvan, M., Taherian T., & Yazdanmehr, E. (2021). Foreign language enjoyment: a longitudinal confirmatory factor analysis–curve of factors model. *Journal of Multilingual and Multicultural Development,* Epub ahead of print. DOI:10.1080/01434632.2021.1874392

Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation.* London: Edward Arnold.

Gardner, R. C. (2019). Looking back and looking forward. In A. Al-Hoorie & P. D. MacIntyre (Eds.), *Contemporary language motivation theory: 60 years since Gardner and Lambert (1959)* (pp. 5–14). Bristol: Multilingual Matters.

Goris, J., Brass, M., Cambier, C., Delplanque, J., Wiersema, J.R., & Braem, S. (2020). The relation between preference for predictability and autistic traits. *Autism Research*, *13*, 1144-1154.

Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 31–50). Bristol: Multilingual Matters.

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

Jiang, Y. (2020). An investigation of the effect of teacher on Chinese university students' foreign language enjoyment. *Foreign Language World*, *196*(1), 60–68.

Jiang, Y., & Dewaele, J.-M. (2019). How unique is the Foreign Language Classroom Enjoyment and Anxiety of Chinese EFL Learners? *System*, *82*, 13-25.

Kiesler, S. B. (1973). Preference for predictability or unpredictability as a mediator of reactions to norm violations. *Journal of Personality and Social Psychology*, *27*(3), 354–359.

Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R.* New York: Routledge

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.

Li, C. (2021). A Control–Value Theory approach to boredom in English classes among university students in China. *The Modern Language Journal*, *105*(1), 317-334.

Li, C., Huang, J., & Li, B. (2020). The predictive effects of classroom environment and trait emotional intelligence on Foreign Language Enjoyment and Anxiety. *System*, *96*, 102393.

Li, C., & Xu, J. (2019). Trait Emotional Intelligence and Classroom Emotions: A Positive Psychology Investigation and Intervention Among Chinese EFL Learners. *Frontiers in Psychology*, *10*, 2453.

Macintyre, P.D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching*, *7*(1), 61-88.

Masgoret, A.-M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and Associates. *Language Learning*, *53*(1), 123–163.

Neff, P., & Rucynski, J. (2017). Japanese perceptions of humor in the English language classroom. *Humor*, *30*(3), 279-301.

Neff, P., & Rucynski, J. (2021). The English classroom as "warai no ba": Instructor views on humor and language learning in Japan. *International Journal of Educational Research Open*, *2*(2). DOI:10.1016/j.ijedro.2021.100070

Pan, C., & Zhang, X. (2021). A longitudinal study of foreign language anxiety and enjoyment. *Language Teaching Research*. Epub ahead of print. DOI:10.1177/1362168821993341

Pavelescu, L. M. (2019). Motivation and emotion in the EFL learning experience of Romanian adolescent students: Two contrasting cases. *Studies in Second Language Learning and Teaching, 9*(1), 55-82.

Saito, K., Dewaele, J.-M., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience and second language comprehensibility development in classroom settings. *Language Learning*, *68*, 709-743.

Talebzadeh, N., Elahi Shirvan, M., & Khajavy, G. H. (2020). Dynamics and mechanisms of foreign language enjoyment contagion. *Innovation in Language Learning and Teaching*, *14*(5), 399-420.

Teimouri, Y. (2017). L2 selves, emotions, and motivated behaviors. *Studies in Second Language Acquisition*, *39*(4), 681-709.

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 489-489.

Tennant, J., & Gardner, R.C. (2004). The computerized mini-AMTB. *Calico Journal, 21*(2), 245-263.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*(5), 2020–2045.

Zhang, H., Dai, Y., & Wang, Y. (2020). Motivation and second foreign language proficiency: The mediating role of Foreign Language Enjoyment. *Sustainability*, *12*(4), 1302.

Zhang, L., & Tsung, L. (2021). Learning Chinese as a second language in China: Positive emotions and enjoyment. *System*, 96, DOI:1010.1016/j.system.2020.102410.

Appendix

Descriptive Statistics of FLE, FLCA, AM, and teacher behavior variables at Time 1, Time 2, Time 3 and Time 4

	Time 1		Time 2		Time 3		Time 4	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Learner variables								
Foreign Language Enjoyment	4.02	0.57	3.92	0.66	3.92	0.71	3.90	0.77
Foreign Language Classroom Anxiety	2.73	0.72	2.75	0.70	2.76	0.74	2.73	0.80
Attitudes/Motivation	5.62	1.07	5.45	1.27	5.34	1.35	5.44	1.35
Teacher behaviour variables								
Foreign language use	6.42	1.05	6.01	1.42	5.83	1.46	5.84	1.47
Predictability	5.49	1.36	5.34	1.50	5.17	1.55	5.24	1.52
Jokes	5.03	1.55	4.77	1.56	4.77	1.57	4.94	1.51

---

[1] Gardner (1985) defined an individual's attitude as "an evaluative reaction to some referent or attitude object, inferred on the basis of the individual belief's or opinions about the referent" (p. 9). He defined motivation as "the combination of effort plus desire to achieve the goal of learning the language plus favorable attitudes towards learning the language (p. 10).
[2] Available at https://publish.uwo.ca/~gardner/docs/englishamtb.pdf
[3] IRB approval number: 173695