Do Girls Have All the Fun?
Anxiety and Enjoyment in the Foreign Language Classroom

Abstract: The present study focuses on gender differences in Foreign Language Enjoyment (FLE) and Foreign Language Classroom Anxiety (FLCA) among 1736 FL learners (1287 females, 449 males) from around the world. We used 21 items, rated on a Likert scale, reflecting various aspects of FLE (Dewaele & MacIntyre, 2014), and 8 items extracted from the FLCAS (Horwitz et al., 1986). An open question on FLE also provided us with narrative data. Previous research on the database, relying on an average measure of FLE and FLCA (Dewaele & MacIntyre, 2014) revealed significant gender differences. The present study looks at gender differences in FLE and FLCA at item level.

Independent t-tests revealed that female participants reported having significantly more fun in the FL class, where they felt that they were learning interesting things, and they were prouder than male peers of their FL performance. However, female participants also experienced significantly more (mild) FLCA: they worried significantly more than male peers about their mistakes and were less confident in using the FL. Our female participants thus reported experiencing both more positive and more mild negative emotions in the FL classroom. We argue that this heightened emotionality benefits the acquisition and use of the FL.

Keywords: Foreign Language Enjoyment, Foreign Language Classroom Anxiety, individual differences, gender
Introduction

Perhaps the most common way to divide humans is into two sexes, male and female. In childhood play, sex differences establish themselves and same-sex friendship lasts a lifetime for many of us, even if we marry a member of the opposite sex. Popular books tell us that “Men are from Mars and Women are from Venus” (Gray, 1992)—so distinct are men and women that they seem to come from different planets. However, not everybody agrees with this perspective. Feminist scholars, such as Cameron (2008), have accused Gray of perpetuating myths and of exploiting “people’s tendency to rely on stereotypes when processing information” (p. 14). Moreover, we acknowledge that gender is not a clear-cut binary category. Transgender people situate themselves at different points along the gender continuum and this can be dynamic, as they can live in role part time (Kulick, 1998).

We agree that it is crucial to avoid simple, stereotypical concepts of gender differences. Eckert and McConnell-Ginet (2013) have shown that the connections between language and gender are deep yet fluid, and arise in social practice. Before speculating on possible reasons for differences between women and men (or the absence of them), there is reason to investigate how large the differences between boys and girls, men and women really are, especially when it comes to language learning. Simply believing the stereotype that girls have superior language learning skill might discourage some boys from engaging with the language learning process, in a self-fulfilling prophesy. The belief in gender differences, and what one should do about them, might be more relevant to communication than the observable gender differences themselves. For this reason, gender differences merit thorough investigation. Such research is particularly relevant in relation to emotional dimensions in FL learning, where small but significant differences have been identified (Dewaele & MacIntyre, 2014).

Literature Review

Meta-analyses point to the discovery of only relatively small gender differences\(^1\) in various behaviors relevant to language, emotion, and communication. What is typically considered female communication style tends to be grounded in use of emotional expression (Tenenbaum, Ford & Alkhedairy, 2011). Stereotypically, women are expected to react to negative events more

---

\(^1\) We use “gender” to refer to the social construct.
emotionally, for example with sadness, crying, and greater withdrawal than men (Hess et al., 2000). This leads to a question as to whether these judgments are due to unsubstantiated gender labels, or whether they have some basis in the intensity of felt emotion.

Neuroscience research shows that significant differences have been found in the way men and women process the same events. Event related potential (ERP) studies have found higher emotional reactivity in women than men when viewing a negative stimulus (Gardener et al., 2013). Furthermore, a comprehensive review by Chaplin and Aldao (2013) found that although women showed more internalizing emotions, such as sadness, anxiety, and fear, they also showed more positive emotions, such as happiness. This may be due to a higher level of emotional reactivity in women overall, not only to negative events.

While stereotypes paint a broad picture of gender differences in language, communication and emotional experience, the meta-analytic work does not support sweeping generalizations. A number of studies have addressed the issue of gender differences in FL learning that follow a trend of statistically significant but small effect sizes. There are studies that show a female advantage in language development and usage. Eriksson et al. (2012) analyzed emerging language skills in ten separate linguistic communities. They found that females show a small yet consistent advantage over males in several aptitudes, including communicative gestures, productive vocabulary, and combining words. Studies of scholastic achievement show similar results. One meta-analysis based on 502 effect sizes indicated that girls had a consistent advantage over boys in school marks, with effect sizes being largest in language-based courses (Voyer & Voyer, 2014). Further, a longitudinal study conducted by Van de Gaer, Pustjens, Van Damme and De Munter (2009) on a cohort of 2,270 secondary school students (aged 12–18) in Flanders (the Dutch-speaking part of Belgium) showed that girls made a quasi-linear positive learning gain in language across their secondary education, whereas boys started with a decline followed by acceleration in their learning gain in language.

Denies (2015) found a more mixed picture in her investigation of the English L2 self-efficacy beliefs of over 22,500 European fifteen-year-olds. She compared listening, reading and writing proficiency to their endorsement of 12 Can-Do statements taken from the Common European Framework of Reference for Languages (CEFR). Logistic regression analyses revealed positive correlations between both elements and variation linked to gender and country of origin. Boys tended to have higher self-efficacy beliefs than equally proficient girls, but the pattern was reversed when students were presented with the easiest task descriptions. Denies (2015) carried out a second study, linking gender to students’ proficiency across three languages, three skills and fourteen countries. English turned out to be the exception to the general tendency for girls to surpass boys’ L2 skills, with boys sometimes performing significantly better than
girls. Gender differences were stronger in writing than in listening or reading. Cross-country variation in the gender gap was interpreted by Denies as evidence that gender is primarily a social factor rather than a biological factor in FL learning. Her multilevel mediation analyses with constituents of motivation suggested that in most cases where girls outperformed boys, a significant part of this advantage could be explained by the differential appeal of the students’ L2 course and by the instrumental, integrative and intrinsic value that students attribute to the L2 (Denies, 2015).

Henry and Cliffordson (2013) noted that gender differences in language motivation research have received little attention (Dörnyei & Csizér, 2002). Existing studies have found that adolescent male FL learners in the United Kingdom (Williams, Burden & Lanvers, 2002), Canada (Kissau, 2006; Kissau & Turnbull, 2008) and Australia (Carr & Pauwels, 2006) have less desire to acquire FLs and invest less effort in learning them than their female peers. Williams et al. (2002) also reported strong gender differences among British learners in their perceptions of FLs: French was the language preferred by girls while boys preferred German. A 14-year-old boy explained that “‘French is the language of love and stuff’” (p. 520) while a 14-year-old girl said that German reflects “‘the war, Hitler and all that’” (p. 520). Kissau and Turnbull (2008) confirmed that for boys, not all languages are perceived equally.

Anxiety certainly has been widely examined in language learning and might be the most studied emotion in Second Language Acquisition (SLA) (Horwitz, 2010). Defined as “the worry and negative emotional reaction aroused when learning or using a second language” (MacIntyre, 1999, p. 27), anxiety-arousal can weaken the learning potential of the FL learner, wreck the best teaching techniques and render the most attractive teaching material inadequate (Arnold & Brown, 1999, p. 2). It has been described as one of the strongest predictors of success or failure in FL learning (MacIntyre, 1999). The list of potential sources of anxiety in language learning and communication is long, including harsh error correction (Gregersen, 2003), competitiveness among learners (Bailey, 1983), incompatibility between teacher and student (Gregersen & MacIntyre, 2014), personality traits such as neuroticism, extraversion, and psychoticism (Dewaele, 2002; 2013a), emotional stability and social initiative (Dewaele & Al Saraj, 2015), perfectionism (Dewaele, to appear; Gregersen & Horwitz, 2002), second language tolerance of ambiguity (Dewaele & Shan Ip, 2013), and other factors (Horwitz, 2010).

Conflicting findings have emerged from research into the relationship between gender and FLA. Some researchers reported lower FLA levels for women than for men (Campbell & Shaw, 1994; Kitano, 2001); others have found the opposite pattern (Arnaiz & Guillén, 2012; Donovan & MacIntyre, 2005) and some have found no differences at all (Dewaele, Petrides & Furnham, 2008, Dewaele, 2013b; Matsuda & Gobel, 2004). Wang (2010) observed that confu-
sion reigns in this area of research and that “unknown variables” may be at play (p. 96). Park and French (2013) attributed the inconsistent research results for gender differences to socio-cultural factors because prior research studies were conducted in one location, and speculated that learner variables, such as motivation, interest and sincerity, might be relevant to the pattern of gender differences. Perhaps in considering the research methods to investigate gender differences, the meta-analytic research net is being cast too wide to capture the nature of the effect—gender differences might not be found on an inter-planetary level but rather on a detail-oriented, microscopic level. In the present study we will examine in detail the possibility that the different findings might be the result of a lack of focus on the details of the emotional reactions themselves, where specific facets of anxiety might be tied to gender differences. Park and French (2013) investigated specific items of the FLCAS, but the authors did not focus on the substantive meaning of those items which might reveal some of the more subtle processes tied to gender.

Although anxiety has been well studied, it is difficult to draw conclusions about gender differences in enjoyment from the existing SLA literature. Previous literature on emotion in language learning has largely concentrated on negative emotion, with positive emotion not being as well researched (Arnold, 2011; Arnold & Brown, 1999; Brown & White, 2010; Imai, 2010). However, positive attitudes and emotions feature prominently in models of a broader complex of motivation. For example, Gardner’s influential model of integrative motivation (1985; 2010) described the motivated learner as having affective reactions that include interest in FLs and desire to learn the FL.

Further study of positive emotions in language learning, as topics in their own right, is warranted by developments in Positive Psychology and potential applications to SLA (see Lake, 2013; MacIntyre & Mercer, 2014). Fredrickson’s broaden-and-build theory has accumulated evidence for a clear differentiation between positive and negative emotions (Fredrickson, 2001; 2003; 2007). Fredrickson has reported that negative emotions tend to be associated with a specific action tendency, that is, a tendency toward a specific type of behavior (see Reeve, 2005). For example, anger leads to the urge to destroy obstacles in one’s path. Fredrickson’s research proposes that, compared with negative emotions, positive emotions produce a qualitatively different type of response.

[The Broaden and Build] theory states that certain discrete positive emotions—including joy, interest, contentment, pride, and love—although phenomenologically distinct, all share the ability to broaden people’s momentary thought-action repertoires and build their enduring personal resources, ranging from physical and intellectual resources to social and psychological resources. (Fredrickson, 2003, p. 219)
The effects of positive emotion go beyond pleasant feelings by enhancing the ability to notice things in the environment, enhancing awareness of language input, and dissipating the lingering effects of negative arousal, promoting personal resiliency and hardiness during tough times. Positive emotion facilitates both exploration and play, two key factors that tend to bring people together.

The development of theories in psychology that differentiate positive from negative emotion raises the interesting question of how each affects language learning (MacIntyre & Gregersen, 2012). In a preliminary investigation, Dewaele and MacIntyre (2014) investigated the amount of overlap between FLE and FLCA. They developed a FLE scale with 21 items (e.g., creativity, pride, interest, fun) reflecting a positive environment in the FL class (teacher and peers). They also used 8 items extracted from the well-established FLCAS reflecting physical symptoms of anxiety, nervousness and lack of confidence (Horwitz et al., 1986). They found a small, negative correlation between FLE and FLCA ($r = -0.36$), and that overall levels of FLE were significantly higher than those of FLCA. Interestingly, female participants scored higher for both FLE ($p < 0.002$) and for FLCA ($p < 0.004$) (p. 254). Dewaele and MacIntyre (2014) thus claim that FLE and FLCA are empirically (negatively) related but are not opposite ends of the same experience—enjoyment and anxiety are distinct emotions.

We decided that a greater amount of detail was needed on FLE and FLCA to understand more precisely what it is that women fear and enjoy more in the FL class. In considering the questions that will guide research we will look at some of the subtle ways in which anxiety and enjoyment operate in the language learning context, and in particular for ways in which male and female learners differ with respect to those emotions. In designing the following research questions, we are keeping the positive emotion (enjoyment) strictly separate from the negative emotion (anxiety) by focusing on the meaning of specific items.

**Research questions**

1. Are there gender differences in items reflecting FLE and FLCA? To be more precise, what items specifically reflect gender differences and what does the wording of particular items tell us about gender and language learning?
2. Given the gender differences, how do women and men describe their FL classroom experiences of anxiety and enjoyment?
Method

Participants and demographics. A total of 1736 multilinguals (1287 females, 449 males) filled out the online questionnaire. The age range of participants was quite broad, ranging from 11 to 75, with a mean age of 24 (SD = 8.5). Levels of education represented in the sample included: having—or working towards—an intermediate high school diploma (N = 91), a high school diploma (N = 113), a Bachelor’s degree (N = 994), a Master’s degree (N = 450) or a PhD (N = 94).

The largest number of participants were studying English as a FL (N = 855, 49%), while others were studying French (N = 276, 16%), Spanish (N = 218, 13%), Dutch (N = 157, 9%) and German (N = 139, 8%), or a range of different languages (6%). A small number of respondents (N = 34) were not currently involved in FL learning, but had had recent FL experience. Most participants (N = 1322, 84%) were studying one FL, with smaller numbers studying two (N = 215, 14%) or three (N = 37, 2%) FLs. A quarter of participants had been studying a FL for 5 years or less, another quarter had studied one between 6 and 9 years, a third quarter had studied a FL between 10 and 12 years, with the remaining participants having studied a FL for more than 13 years.

When asked about their mastery of the FL overall, only 5 described themselves as beginners (0.3%), 243 as low intermediate (14%), 67 as intermediate (4%), 435 as high intermediate (25%) and 293 as advanced (17%).

Asked about how they would rate themselves compared to their FL learning peers, 42 described themselves as far below average (2%), 171 as below average (10%), 772 as average (44%), 638 as above average (37%), and 116 as far above average (7%).

Of the 90 nationalities reported, Belgians constituted the largest group (N = 365, 21%), followed by Britons (N = 244, 14%), Chinese (N = 174, 10%), Americans (N = 118, 7%), and 845 participants (48%) belonging to other nationality groups (N < 80). There were also many participants who self-reported as having dual-nationality.

The sample consists of 456 self-reported bilinguals (26%), 555 trilinguals (32%), 415 quadrilinguals (24%), 202 pentalinguals (12%), and 70 sextalinguals (4%). The 44 remaining participants (3%) reported knowledge of 7 to 10 languages.

---

2 10 participants did not disclose their gender.
3 172 participants did not provide this information.
4 We decided to merge the category of beginners with that of lower-intermediate.
The instrument. The questionnaire started with a demographics section from which the above information was retrieved. Following this, participants were asked to respond to 29 items describing the FL class on standard 5-point Likert scales with the anchors “absolutely disagree” = 1, “disagree” = 2, “neither agree nor disagree” = 3, “agree” = 4, “strongly agree” = 5. Of these items, 8 were extracted from the FLCAS (Horwitz et al., 1986) and 21 items written to reflect FLE (Dewaele & MacIntyre, 2014). The starting point for the FL enjoyment scale was Ryan et al.’s (1990) Interest/Enjoyment sub-scale consisting of 7 items related to enjoyment, fun, interest, and boredom at one point in time. These were specifically adapted to the FL environment and were rephrased so that they elicited a more global judgment of past FL classes. We added “items relating to dealing with FL mistakes made in public, identity, improvement in using the FL, pride in one’s own performance, group membership, the social environment and cohesiveness, attitudes towards the learning of the FL, the presence of laughter, and judgments about peers and teachers” (p. 243). All items were positively phrased. In addition, the 8 FLCAS items reflected physical symptoms of anxiety, nervousness, and lack of confidence (Horwitz et al., 1986). They were chosen to capture the reliability of the original scale without sacrificing the reliability of the measurement (MacIntyre, 1992). Two FLCA items were phrased to indicate low anxiety and six were phrased to indicate high anxiety. The low anxiety items were reverse-coded so that high scores reflect high anxiety for all items on this measure. A one-sample Kolmorogov-Smirnov test revealed that the distribution of the items was not normal. However, t-tests tolerate moderate violations to their normality assumption (Rosenkrantz, 2008, p. 478) and the Levene’s Test for Equality of Variances will be used to verify the assumption of equality in the groups of females and males.

The final survey question was open-ended. The instructions were as follows: “Describe one specific event or episode in your FL class that you really enjoyed, and describe your feeling in as much detail as possible.” One thousand and seventy-six participants answered the open question, producing 52,471 words. These data are used for illustrative purposes only.

The first version of the questionnaire was pilot-tested with 15 participants. This led to the deletion of some items and the reformulation of others. The research design and questionnaire obtained approval from the Ethics Committee of the School of Social Sciences, History and Politics at Birkbeck, University of London. The final version of the questionnaire was posted online using Googledocs and an open call was addressed to multilinguals, to colleagues and friends of the authors, including participants in previous studies, asking them to forward the call to friends, teachers or students. In other words, we used snowball sampling, that is, a non-probability sampling technique. We are aware that there was an inevitable self-selection bias, as it is more likely for
satisfied language learners to participate in a study on language learning than those who dislike it.

Data analysis. Our data analysis can be described as the data-validation variant of a convergent parallel design (Cresswell & Plano Clark, 2011). In this approach, quantitative and qualitative data are collected simultaneously, which is efficient with large samples. The purpose of the quantitative data in this approach is to identify the specific items where values of male and female participants differed significantly. The purpose of the qualitative data is to provide illustrations of the closed items in the questionnaire, written by participants themselves. Results provided by the two types of data are mixed and combined in the discussion section. Cresswell and Plano Clark (2011, p. 81) explain that with this type of analytic approach “...the qualitative items are an add-on to a quantitative instrument, the items generally do not result in a complete context-based qualitative data set. However, they provide the researcher with emergent themes and interesting quotes that can be used to validate and embellish the quantitative survey findings.”

The analysis will proceed in two stages. In the first phase, the statistical analysis will be used to identify which of the items had significantly different values for the male and female participants. In the second phase, data extracts will be selected from the open question guided by the principle that they are representative of a particular closed item, concise, and interesting. In this way, we give a voice to participants, allowing us to add an emic dimension to the abstract, etic statistical dimensions.

Results

In order to assess gender differences in the FLE and FLCA items, a series of independent *t*-tests were performed. In reporting the results below, we adjusted the probability level required to decide on statistical significance using Bonferroni’s correction to reduce the risk of type I error. The *t*-test is considered to be a relatively liberal test, and the Bonferroni adjustment is fairly conservative. Adopting this procedure might produce a slightly elevated risk of Type II errors overall, that is, we might not be declaring a difference to be significant when it should be. For this reason, we make note of both the Bonferroni-adjusted and unadjusted significance levels.

For this set of analyses, the *p*-value required for significance is equal to or smaller than 0.002 (0.05/29 = 0.002). A series of independent *t*-tests revealed that female participants differed significantly from males for 5 out of 29 to-
tal items, with significant differences for another 8 items using the standard $p < 0.05$ criterion (see Table 1).

Table 1

<table>
<thead>
<tr>
<th>Item</th>
<th>$t$</th>
<th>$df$</th>
<th>$p$</th>
<th>Mean females</th>
<th>Mean males</th>
</tr>
</thead>
<tbody>
<tr>
<td>In class, I feel proud of my accomplishments</td>
<td>4.1</td>
<td>771.4</td>
<td>0.001</td>
<td>3.84</td>
<td>3.65</td>
</tr>
<tr>
<td>I don't worry about making mistakes in FL class</td>
<td>3.9</td>
<td>1734</td>
<td>0.001</td>
<td>3.05</td>
<td>3.30</td>
</tr>
<tr>
<td>It's fun</td>
<td>3.4</td>
<td>1734</td>
<td>0.001</td>
<td>4.34</td>
<td>4.18</td>
</tr>
<tr>
<td>I feel confident when I speak in FL class</td>
<td>3.1</td>
<td>1734</td>
<td>0.002</td>
<td>3.13</td>
<td>3.31</td>
</tr>
<tr>
<td>I've learnt interesting things</td>
<td>3.1</td>
<td>704.5</td>
<td>0.002</td>
<td>4.32</td>
<td>4.18</td>
</tr>
<tr>
<td>I can feel my heart pounding when I'm going to be called on in FL class</td>
<td>3.0</td>
<td>1734</td>
<td>0.003</td>
<td>2.99</td>
<td>2.78</td>
</tr>
<tr>
<td>I get nervous and confused when I am speaking in my FL class</td>
<td>2.8</td>
<td>1734</td>
<td>0.005</td>
<td>2.78</td>
<td>2.61</td>
</tr>
<tr>
<td>I enjoy it</td>
<td>2.7</td>
<td>1734</td>
<td>0.006</td>
<td>4.12</td>
<td>3.99</td>
</tr>
<tr>
<td>It's cool to know a FL</td>
<td>-2.1</td>
<td>653</td>
<td>0.011</td>
<td>4.63</td>
<td>4.53</td>
</tr>
<tr>
<td>I always feel that the other students speak the FL better than I do</td>
<td>2.3</td>
<td>1734</td>
<td>0.022</td>
<td>2.82</td>
<td>2.67</td>
</tr>
<tr>
<td>I can be creative</td>
<td>2.2</td>
<td>1734</td>
<td>0.031</td>
<td>3.77</td>
<td>3.67</td>
</tr>
<tr>
<td>I don't get bored</td>
<td>2.1</td>
<td>1734</td>
<td>0.033</td>
<td>3.64</td>
<td>3.51</td>
</tr>
<tr>
<td>There is a good atmosphere</td>
<td>2.0</td>
<td>1734</td>
<td>0.045</td>
<td>4.10</td>
<td>4.01</td>
</tr>
<tr>
<td>Making errors is part of the learning process</td>
<td>-1.7</td>
<td>1734</td>
<td>0.087</td>
<td>4.46</td>
<td>4.52</td>
</tr>
<tr>
<td>I'm a worthy member of the FL class</td>
<td>1.6</td>
<td>739.7</td>
<td>0.089</td>
<td>3.64</td>
<td>3.56</td>
</tr>
<tr>
<td>It's a positive environment</td>
<td>1.5</td>
<td>669.2</td>
<td>0.127</td>
<td>3.99</td>
<td>3.92</td>
</tr>
<tr>
<td>I can laugh off embarrassing mistakes in the FL</td>
<td>1.4</td>
<td>719.2</td>
<td>0.142</td>
<td>3.47</td>
<td>3.39</td>
</tr>
<tr>
<td>We laugh a lot</td>
<td>1.3</td>
<td>1734</td>
<td>0.154</td>
<td>3.45</td>
<td>3.37</td>
</tr>
<tr>
<td>I learnt to express myself better in the FL class</td>
<td>-1.1</td>
<td>1734</td>
<td>0.265</td>
<td>3.38</td>
<td>3.44</td>
</tr>
<tr>
<td>I feel as though I'm a different person during the FL class</td>
<td>1.0</td>
<td>1600</td>
<td>0.290</td>
<td>2.76</td>
<td>2.69</td>
</tr>
<tr>
<td>I start to panic when I have to speak without preparation in the FL class</td>
<td>1.0</td>
<td>1734</td>
<td>0.297</td>
<td>2.74</td>
<td>2.67</td>
</tr>
<tr>
<td>Even if I am well prepared for the FL class, I feel anxious about it</td>
<td>-1.0</td>
<td>811.1</td>
<td>0.324</td>
<td>2.64</td>
<td>2.58</td>
</tr>
<tr>
<td>We form a tight group</td>
<td>0.6</td>
<td>1734</td>
<td>0.518</td>
<td>3.28</td>
<td>3.24</td>
</tr>
<tr>
<td>The teacher is friendly</td>
<td>-0.6</td>
<td>1734</td>
<td>0.533</td>
<td>4.25</td>
<td>4.28</td>
</tr>
<tr>
<td>The teacher is encouraging</td>
<td>-0.5</td>
<td>1734</td>
<td>0.600</td>
<td>4.18</td>
<td>4.21</td>
</tr>
<tr>
<td>We have common &quot;legends,&quot; such as running jokes</td>
<td>0.5</td>
<td>829.4</td>
<td>0.642</td>
<td>3.02</td>
<td>2.99</td>
</tr>
<tr>
<td>The peers are nice</td>
<td>-0.4</td>
<td>1734</td>
<td>0.644</td>
<td>4.01</td>
<td>4.02</td>
</tr>
<tr>
<td>It embarrasses me to volunteer answers in my FL class</td>
<td>0.3</td>
<td>1734</td>
<td>0.774</td>
<td>2.53</td>
<td>2.51</td>
</tr>
<tr>
<td>The teacher is supportive</td>
<td>0.1</td>
<td>1734</td>
<td>0.903</td>
<td>4.20</td>
<td>4.20</td>
</tr>
</tbody>
</table>
From Table 1 we can conclude that the female participants reported stronger negative and positive emotions than the male participants on items that showed gender significant differences.

Figure 1 visualizes the differences between the means of the female and male participants for the five significant items ($p < 0.002$).

In terms of FLE items, using the Bonferroni adjusted significance test and compared to the male participants, the female participants felt significantly more proud of their achievement, had more fun, and felt they were learning interesting things. Using the unadjusted significance test, the females also tended to like the atmosphere of the FL class more, enjoyed it more and agreed more strongly that it is “cool” to know a FL. Finally, females also tended not to feel bored and felt more creative in the FL class, compared to males.

To understand the results in more detail, and to hear the learners’ voices, we examined the comments written by the respondents. The following comments by five female participants illustrate the ways in which emotional reactions unfold, in their own words:

Brittany, female, 18, Canadian: It was the beginning of the school year and the teacher was handing back an essay we wrote the week before and when she gave me mine, she announced to the class that I had the highest mark and that she has never given anyone a mark that high before. I felt really proud of myself because the week before I got back an assignment I thought I did well on but I actually did really bad and it made me start to question if I continuing to learn the French language was the right decision for me. But getting such a good mark on that essay renewed my faith in
myself and it made me determined to always do my best and to not give up when it gets tough.

Sophie, female, 12, British: It is really fun as you make new friends with people you did not know and you form relationships over the bonds of you learning and making mistakes together.

Rosalind, female, 17, British: I had a conversation with someone entirely in French and saw how exciting (and useful) it is to be able to speak in a completely different language. I could also see how all my lessons weren’t a waste of time and that I knew a lot more French than I thought I did.

Amy, female, 17, British: When I was training for my AS Speaking Exams last year, we had a language assistant over from South America to help us. In one of these sessions, I had a light bulb moment. I was suddenly able to express myself in Spanish in a way I hadn’t previously been able to. I was answering spontaneous questions with ease and feeling confident enough to answer without pausing for more than a second.

Inès, female, 19, Belgian: When we played a game that consisted in finding the title of the film describes. It was funny and interesting. It was the first time I really enjoyed the course. And this subject was motivating, I had more things to say and more creativity.

In terms of FLCA items, using the Bonferroni-adjusted significance test, we found that female participants were significantly more worried about their mistakes and significantly less confident than their male peers. They tended to feel more nervous and confused, experienced more physical symptoms of anxiety such as a pounding heart in the FL class and felt the others were better at speaking the FL than they (according to the unadjusted significance tests).

The following observations by three female participants show how the lack of confidence and the worry about mistakes can be overcome by specific activities and by a positive atmosphere in the class:

Caela, female, 22, American: Irish Gaelic has some pretty nice curses (e.g., “May you be eaten by a cat, and may the Devil eat the cat!”), so the day we learned those we practiced cursing each other. We were a small class of about 5 people, so it was really fun to be a little silly. I didn’t feel quite as intimidated to perform well when everyone was laughing, and was able to relax more than usual and try phrases without worrying about having perfect pronunciation.

Audrey, female, 24, Belgian: In English, I like the works of group and when we speak about topical subject. One day with my professor of English we played small plays. At the beginning I was anxious not to manage
to pronounce well the text or to have made the errors in the text which I had written. But at the end I was satisfied to have been able to surmount (overcome) my anxiety to speak in front of everybody in English. I am also satisfied when I receive from good grade after an examination when I know that I worked hard.

Simona, female, 22, Macedonian: We were supposed to have a 2-minute speech before our peers and our professor on a topic we chose. All of my peer’s presentations were great and I really loved experiencing the feeling when I was talking before all of them. It was really special, even if at first I was a bit nervous and felt my heart pounding, but it felt great standing there and expressing my opinion and knowing that all of the other students are listening to you with attention.

Male participants made similar comments. Philip reported the joy of getting the pronunciation right:

Philip, male, 23, Swedish: When I nailed the pronunciation on a sentence I read out loud in my seminar group.

Cristian enjoyed a moment of laughter in class that stimulated his learning:

Cristian, male, 18, Columbian: El momento que me gusto fue cuando jugamos domino esto fue un momento de risa en el cual me diverti aprendí y fue un momento muy agradable. (‘The moment I enjoyed was when we played domino, it was a moment of laughing, it was fun, I learned and it was very nice’).

No significant differences appeared for three FLCA items referring to the more paralyzing aspects of anxiety (‘I start to panic when I have to speak without preparation in the FL class’; ‘Even if I am well prepared for the FL class, I feel anxious about it’; ‘It embarrasses me to volunteer answers in my FL class’). Attitudes towards errors also were similar in both groups (‘Making errors is part of the learning process’; ‘I can laugh off embarrassing mistakes in the FL’). The same was true for perceived progress (‘I learnt to express myself better in the FL class’). There were no differences either in the judgment of the social aspects in the FL class (‘It’s a positive environment’; ‘We form a tight group’; ‘We have common legends such as running jokes’; ‘The peers are nice’), nor in the appreciation of the teachers (friendly, encouraging, supportive). Finally, no differences emerged between both groups in the amount of laughing, in the ability to express oneself in the FL, in feeling a worthy member of the group and in assuming a different identity in the FL class.
Discussion

The present study reveals a rich and nuanced picture of gender differences in two emotional dimensions, anxiety and enjoyment. We knew from Dewaele and MacIntyre (2014) that female participants scored significantly higher on the composite measures of FLE and FLCA. Although female participants experienced more FLCA overall, gender differences were not significant on all items. A similar pattern emerged in Park and French (2013) who reported gender differences on only 5 of the 33 items on the complete FLCAS scale. In the present data, the difference between female and male respondents was only significant for two items reflecting relatively milder experiences of FLCA. Female participants worried more about mistakes and were less confident. They also tended to feel that the other students were better, they tended to feel more nervous and confused, and were more likely to have a pounding heart in the FL class. However, they did not differ from the male participants in the more severe, paralyzing aspects of anxiety, such as panic, freezing-up often out of embarrassment, or increased anxiety despite being prepared. Williams et al. (2002) reported similar results, with girls scoring higher on desire and liking the FL, but not showing a significant difference with the boys in their attitudes towards teachers, parents and peers.

Our female participants had more fun, pride, enjoyment, excitement, and interest in the FL class that allowed them to be creative and become “cool” multilinguals. The observations of Brittany, Sophie, Rosalind, Inès, Amy, Philip and Cristian reported above bring the statistical findings to life. In their own words, learners described how a good mark on an essay boosted their pride and strengthened their determination to do well in the FL class, how learning together with peers is fun because it strengthens social bonds, how excited one can feel when actually being able to use the FL, and how funny and interesting games in the FL class can allow students to be creative and boost their motivation. These specific contexts converge with Chaplin and Aldao’s (2013) finding that girls show more positive and internalizing emotions (sadness, anxiety, sympathy) than boys, especially in adolescence, a time often associated with emotional ups and downs. This obviously does not mean boys cannot experience similar emotions, as our comments showed. When interpreting gender differences, we should carefully avoid thinking of the results in terms of mutually exclusive categories (as in the Mars versus Venus type of arguments), preferring instead to think of baseline differences between groups that are modified to a considerable extent by individual experiences and lead to a wide distribution of scores within each group, and considerable overlap between groups.

Similar caution is necessary when conceptualizing positive and negative emotions. Although it might seem natural to conceptualize positive and negative emotion as operating in a seesaw fashion (one goes up and the other goes
Do Girls Have All the Fun?...

down), such a view is inconsistent with emotion research, especially over the past decade (Fredrickson, 2007; MacIntyre & Gregersen, 2012). The combination of both higher FLCA and higher FLE shown by female learners is not a contradictory result. Female FL learners seem to care more about their FL performance and their progress in the FL perhaps because they find it intrinsically appealing, fun and of value (Dörnyei & Csizér, 2002; Henry & Cliffordson, 2013). In effect, this pattern “raises the stakes” for learning, boosting both positive and negative emotion. Although there is some worry that they might not do as well in the FL as they expect, the concern does not tend to cause them to panic or freeze when having to use the FL. Respondents Caela, Audrey, and Simona described how they managed to overcome anxiety and perform in the FL in front of their peers and teacher.

Fredrickson (2001; 2007) argues that enjoyment is associated with characteristics most teachers would welcome: creativity, the urge to play, and going beyond one’s limits. Experiencing joy in the FL class is described as a facilitating experience for the language learners in our sample whose prose nicely exemplifies Fredrickson’s theoretical tenets. Fredrickson (2001; 2013) points out that the different types of resources that are built during play, including social bonds and specific skills, last much longer than the transient feelings that set the emotional tone for learning. If female, compared to male, FL learners are generally having more fun in FL classes, they might unlock their potential faster and thus progress further than their male peers, even if they also experience some of the facets of anxiety more often or with greater intensity.

High levels of language anxiety are typically negatively linked to FL proficiency scores, including course grades, standardized tests, and other measures (Horwitz, 2010). Researchers generally agree that strong anxiety feelings are disruptive to behavior, interfering with interpersonal communication, cognition, and learning (MacIntyre & Gardner, 1994). However, the interpretation of cases of moderate anxiety, and what to do about less-severe levels of anxiousness, might change considerably if we consider a wider emotional context. Dewaele and MacIntyre (2014) argued that enjoyment and anxiety could cooperate from time to time, with enjoyment encouraging playful exploration and anxiety generating focus on the need to take specific action. Our data suggests that we find greater levels of enjoyment among the active and successful learners in our sample, yet some anxiety certainly is present. Is it possible that our female participants managed to find a constructive balance between enjoyment and anxiety? Did they manage to develop a better ratio of positive to negative emotion in the FL class (Fredrickson, 2013)?

Females scoring higher than males on FLE and FLCA items is not problematic conceptually or in practice; a combination of stronger positive and negative emotions in FL learning might provide a stronger basis for motivation than
would weaker emotions. It can be argued that both positive and negative emotions are inherently adaptive (MacIntyre, 2007; Reeve, 2005), suggesting that stronger overall emotional experiences can facilitate language learning even if negative emotions are part of the experience. Fredrickson (2013) has suggested that the ratio of positive to negative emotions might be more important than simply the absence of negative emotion. Typically SLA studies of emotions have focused on one specific emotion, most often anxiety. But if we consider emotions in combination with each other, then a study focused on anxiety (e.g., Park & French, 2013) or one that is designed to arouse anxiety (MacIntyre & Gardner, 1994) is likely activating a suite of associated emotions, both positive and negative (such as pride and confidence, frustration and anger) among the participants. Even if the focal point is anxiety, we should be aware of the network of emotions that surround it.

Mixing emotions occurs not only at the item level, but at the research design level as well. Interpreting the mixture of positive and negative emotions raises an important methods issue, specifically aggregation of regular and reversed-scored items. When a scale has some items with positive wording and other items with negative wording, it is the researcher who aggregates positive and negative emotions into a single construct, a single score for each respondent’s FLA for example. This is done by taking the negatively worded items and “flipping” the scores (e.g., a “1” becomes a “7” and a “2” becomes a “6” on a 7-point Likert scale) before adding them to scores on the other items. In the case of FLCAS and other anxiety measures, positively phrased item-level words such as proud, confident, sure, and relaxed are reverse-scored at the time of analysis. But for the research participants, scores on these items might reflect something phenomenologically different as they respond to the questionnaire items. If positive and negative emotions are correlated but essentially different experiences, as Fredrickson suggests, then existing measures of anxiety may be conceptually less homogeneous than generally assumed because of the mix of positive and negative emotion.

It should be emphasized that the reverse-scoring used in self-report scales of emotions can be justified on the basis of item-response theory. It is a threat to the validity of scores when items are all keyed in a positive or negative direction. In addition, the correlations among items and the overall Cronbach’s alpha coefficient typically obtained for the FLCAS (often in excess of 0.90) justifies the procedure of reverse-scoring and aggregation. It is when one takes on a research question with a more microscopic-level analysis, as we did in this study, that the mixing of emotions by using reverse scoring procedures presents itself as a conceptual issue. If the purpose of research is to assess changes in

---

5 This is not the case with models of motivation such as Gardner’s (1985; 2010) integrative motive which has combined anxiety with positive emotions such as interest and desire to learn.
Do Girls Have All the Fun?…

typical anxiety reactions over time, or correlations between FLA and other measures (such as course grades), then reverse-scoring FLCAS items does the job very well (Horwitz, 2010). But if the research purpose is to dissect subtle variations in positive and negative facets of emotions as we are doing in the present study, or to identify changes in emotion over brief timescales, then we must exercise additional caution. Patterns and processes that are observed on one level or timescale can change substantially when examined on either a broader or narrower basis.

Pedagogical implications drawn from this analysis might help settle a long standing issue in the anxiety literature—the nature of facilitating anxiety (Scovel, 1978; Park & French, 2013). Horwitz (1990, in press) has consistently cautioned against language teachers deliberately increasing anxiety among their students in the hope of finding its elusive facilitating effects. This is wise counsel. Our data support and elaborate Horwitz’ admonishment; techniques designed to increase anxiety divorced from enjoyment seem both troublesome and indefensible. However, mild-to-moderate levels of anxiety reactions likely will accompany positive emotions when participants value what they are doing, such as when the self-concept is implicated (e.g., a fun but potentially embarrassing activity) or when something important is at stake (e.g., course grades or future opportunities) (see also Dewaele, 2011; 2015). The subtle mixing of positive and negative emotion at the research design and item aggregation levels potentially has set the stage for the persistent notion of facilitating anxiety. We suggest that it is not the negative anxiety component of the FLCAS but rather the undertones of positive emotion implicated in some of the items that are associated with any potentially facilitating effects of anxiety. Therefore, consistent with the original differentiation of facilitating and debilitating anxiety (Alpert & Haber, 1960), facilitating anxiety should not be considered a lower level or amount of anxiety (below a hypothetical ‘threshold’) but rather a qualitatively different, strongly-negatively correlated, positive emotion related to excitement, risk-taking, arousal, eagerness, and so on. One potential implication of this notion is that teachers who put the prime focus on creating enjoyment experiences for learners in learning situations where something is at stake (e.g., course grades, public performance, and learner self-esteem) likely have to accept that some anxiety might come along for the ride, especially for the female learners. The present data that show that positive and negative emotions indeed are mixing in language classrooms; learners with greater emotional intensity are likely to experience both more anxiety and more enjoyment in FL classrooms.

Before concluding this article we offer some comments on the limitations of the research. Our sample is both very large and mostly self-selected. As such, the respondents do not necessarily represent the general population of FL learners (if such a sample actually could be defined and meaningfully delineated). In addition, participants who volunteered to spend time completing the online
questionnaire probably were positively biased towards FL learning in the first place. The size and international scope of the present sample, though it was not randomly chosen, does help to reduce the concern for gender differences tied to one specific location or nation (Park & French, 2013).

Further research could investigate whether the gender patterns in FLE and FLCA are similar for specific FLs. Williams et al.’s (2002) finding of stereotypical representations about so-called “feminine” and “masculine” languages could affect both FLE and FLCA of female and male FL learners. Here again, caution is needed. Indeed, more girls (69%) than boys studied French up to A level in the UK in 2013, but they also outnumbered them (60%) for German (Report by the Joint Council for Qualifications, 2013). Moreover, as Denies (2015) showed, gender differences can vary from country to country, and from skill to skill. In other words, having a sample from a single geographical and school context might allow researchers to control certain variables, but it would not necessarily increase the generalizability of the findings.

We are aware that dividing our population into male and females learners may lead to the objection that we ignored possible intra-group differences and that gender may combine with contextual factors that influence FLCA and FLE. The present study is relatively decontextualized, as that is the nature of the survey method when a large, international sample is used. Future research on gender differences that takes into account the effect of classroom environment within specific schools is recommended. For example, it would be an interesting research question to ask whether male learners, who have been found to prefer a relaxed and supportive environment compared to female learners, experience different levels of FLE and FLCA when such a classroom environment is provided (Kissau & Salas, 2013). Finally, our selection of more extracts by female learners than male learners does not imply that the latter differed in the type of observations they made.

Conclusion

Item-level statistical analysis of the FLE scale (Dewaele & MacIntyre, 2014, 2016) and items selected from the FLCAS (Horwitz et al., 1986) revealed that

6 With the caveat that only small differences existed in the proportion of British girls and boys obtaining A* to C grades in 2013 for the A level exams for French, German and Spanish. A larger proportion of girls did obtain A* to C grades for Irish, Welsh, and other modern languages. Of course, one could argue that only a minority of the best male FL learners chose a FL as A level subject, which would explain why their scores were not that different from those of their female peers.
our female participants had significantly more fun, were prouder of their FL performance and felt they were learning interesting things in the FL class. They tended to experience more enjoyment and more excitement in a positive FL classroom environment that allowed them to be creative, and tended to agree more that knowing a FL was “cool” compared to the male participants. The female learners’ superior positive emotions were complemented by a higher level of some facets of anxiety. In other words, the female participants reported experiencing more emotion overall in the FL classroom, which has been described as the engine behind linguistic progress (MacIntyre, 2007). At first glance, this might seem to be a contradiction. However, drawing on Fredrickson’s theory allows for a more complete understanding of the network of positive and negative emotions active in the language learning process. These findings point to the need for research to examine gender differences in a specific context of interest (Hyde, 2005).

Acknowledgment. We would like to thank all our participants for having provided us with precious input and observations. Special thanks also to Dr Ruxandra Comanaru and Dr Ren Wei, who managed the online questionnaire on Googleforms and on a Chinese website. We also want to thank the colleagues and friends who forwarded our call for participation, including Dr Françoise Masuy from the Université Catholique de Louvain and Miss Amy Botwright, University of Cambridge, who helped us recruit participants.

References


Ob Mädchen Spaß haben? Angst und Freude im Fremdsprachenunterricht

Zusammenfassung
