BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION

BOĞAZİÇİ ÜNİVERSİTESİ EĞİTİM DERGİSİ

Vol. 41-3(1)
2024

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
BOĞAZİÇİ ÜNİVERSİTESİ EĞİTİM DERGİSİ

ISSN: 2822-5600
e-ISSN: 2822-5597

VOLUME 41-3(1)
2024

CONTENTS

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
2024, Vol. 41-3, No. 1
https://doi.org/10.52597/buje.1389669

The Role of Native Speaker Peers on Language Learners' Fear of Negative Evaluation and Language Anxiety

Tuba Yılmaz[a] and Ester De Jong[b]

Abstract

This study investigated graduate-level international students' linguistic and learning experiences in classrooms predominated by native English-speaking (NES) students. The central objective was to discern the impact of NES students' presence on the fear of negative evaluation experienced by international students speaking English as a second language in an American university. A survey was initially administered to 22 graduate-level international students, and five international students, distinguished by heightened levels of anxiety associated with negative evaluation in classrooms primarily dominated by NES students, were interviewed. The interview data were analyzed with domain analysis. The findings elucidated that international students' fear of negative evaluation depended mainly on instructors' practices and international students' perceptions of NES peers' attitudes. Moreover, the increased fear of negative evaluation adversely impacted their linguistic behaviors and overall engagement with instructional content. The findings implied a need for more inclusive and safe learning environments in graduate schools.

Keywords: fear of negative evaluation, native English speaker, language anxiety, languaging, participation

Article info

Received: 12.10.2023
Revised: 06.12.2023
Accepted: 22.12.2023
Published online: 22.04.2024

Introduction

Language anxiety has gathered profound attention in the last few decades due to its detrimental impacts on English language learners' academic achievement, comprehension and production in the target language, and classroom engagement of students (Aida, 1994; Hewitt & Stephenson, 2012; Horwitz et al., 1986; MacIntyre & Gardner, 1991; Trang et al., 2012; Xianping, 2003). Studies revealed that when second/foreign language learners experienced a high level of language anxiety, they often struggled to concentrate, had heart palpitations or dry throat, felt worried and scared, and even cried during tests due to language anxiety (Cheng, 2014; Horwitz et al., 1986; MacIntyre & Gardner, 1991; Ran et al., 2022; Tanveer, 2007). Moreover, since they perceived 'languaging' as a threatening practice, they often sat in the back of the classroom, refrained from expressing their own opinions, and demonstrated lower school achievement than their native speaker counterparts (Conway, 2007; Pappamihiel, 2002).

One of the components of language anxiety is the fear of negative evaluation. Studies revealed that language learners mostly experience fear of negative evaluation during listening and speaking (Ran et al., 2022) and often prefer staying silent in classrooms dominated by native English-speaking (NES) peers (Morita, 2004; Pappamihiel, 2002). This way, they could avoid making mistakes and receiving unfavorable judgments from their NES peers with stronger language skills (Hanh, 2020; Hamouda, 2013; Price, 1991). To better understand the impact of NES peers' presence on language learners' learning and languaging experiences, this study explored international graduate students' experiences in two different classroom contexts: in classes where the majority of the classroom population was native speakers and in the classes, where the majority of the classroom population was international students. The study has important implications for professors teaching international students in classrooms predominantly populated with NES students.

[a]*Necmettin Erbakan University, Department of Foreign Language Education, tubayilmaz@erbakan.edu.tr, ORCID: 0000-0003-2844-1230*
[b]*University of Colorado Denver, Culturally & Linguistically Diverse Education, ester.dejong@ucdenver.edu, ORCID: 0000-0002-6165-3828*

To cite: Yılmaz, T., & De Jong, E. (2024). The role of native speaker peers on language learners' fear of negative evaluation and language anxiety. *Boğaziçi University Journal of Education, 41-3*(1), 1-17. https://doi.org/10.52597/buje.1389669

Language Anxiety

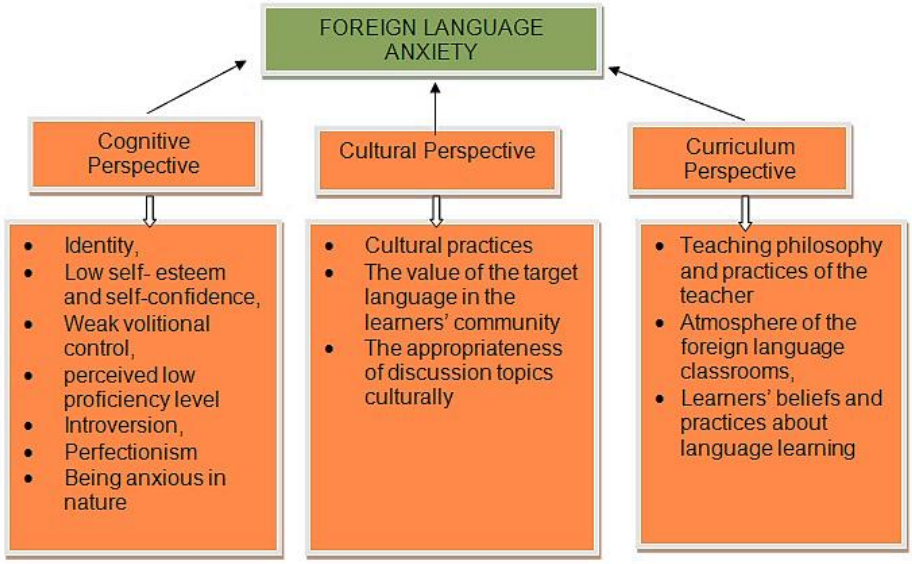
Anxiety includes trait, state, and situation-specific anxiety (Pappamihiel, 2002). While trait anxiety is considered a personal characteristic of individuals, state anxiety appears in the presence of some factors, such as vicarious experiences. Test anxiety, public anxiety, or social anxiety are considered some types of state anxiety. Finally, situation-specific anxieties consider anxiety an emotion that appears in certain contexts.

Language anxiety is "a distinct complex of self-perceptions, beliefs, feelings and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (Horwitz et al., 1991, p. 31). It is considered situation-specific since it appears specifically in second/foreign language-learning contexts due to unique situations that occur exclusively in foreign language courses.

Studies highlighted that multiple factors might elicit language anxiety, and its subtle impacts on the language-learning process can have detrimental outcomes for those with a high degree of language anxiety (Macintyre & Gardner, 1991; Oad et al., 2020; Ohata, 2005; Xianping, 2003; Young, 1991). Zheng (2008) classified these factors into three groups (see Figure 1).

Figure 1

Potential Sources of Language Anxiety (Zheng, 2008)



The cognitive perspective explores the impact of anxiety on the efficient cognitive processing of information in the target language, which, in turn, impairs students' capacity to perform well in class activities and evaluations (MacIntyre & Gardner, 1994; Tobias, 1986; Zheng, 2008). Individual factors such as identity, low self-esteem and self-confidence, weak volitional control, introversion, perfectionism, being naturally anxious, worry, and sentimentality were found to impede the cognitive processing of information (Conway, 2007; Eysenck, 1979; Oxford, 1999; Pappamihiel, 2002; Trang et al., 2012; Zheng, 2008). Studies revealed that language learners with high language anxiety were more likely to be adversely affected by extraneous activities in the classroom and that these distractions had a detrimental impact on their working memory (Eysenck, 1979). Furthermore, due to intrusive cognitive processes, the subjects experienced difficulty maintaining focus during text comprehension. Their vocabulary acquisition exhibited a decelerated pace, accompanied by difficulties retrieving recently learned lexical items. Additionally, the individuals encountered impediments in the timely completion of classroom assignments, as documented by Kasap and Power (2019), Krashen (1985), Maclntyre and Gardner (1989), Lee (2009), Morita (2004), and Sellers (2000).

The cultural perspective analyzes the practices considered normal and appropriate in language learners' cultures but unacceptable and improper in the target culture because these practices can hinder learners' comprehension and classroom participation (Ohata, 2005). The practices, such as taking turns or raising hands to volunteer for answers, can be perceived differently and face-threatening by some second language (L2) learners, provoke their language anxiety levels, and lead to undesirable results. For example, Ohata (2005) reported in his

qualitative study with five Japanese language learners that Japanese cultural values, such as avoiding expressing ideas or not being assertive during discussions, did not align with the target culture. Thus, they increased the learners' anxiety levels and hampered their active classroom participation.

The curriculum perspective examines classroom factors contributing to language anxiety, such as teaching practices, the safety of the learning environment, teachers' error correction strategies, and teachers' attitudes to mistakes (Malik et al., 2020; Tercan & Dikilitas, 2015; Zheng, 2008). These factors can impact language-learning processes, language production, and classroom participation since they can threaten language learners' faces. For example, Fandiño Parra's research (2010) with 17 beginner-level English as a Foreign Language (EFL) students found that learners experienced more anxiety when teachers implemented classroom activities requiring a more spontaneous and authentic use of the foreign language. Moreover, Young (1991) noted that language learners with high language anxieties refrained from languaging in classrooms due to the teacher's excessive correction of errors as it was perceived as offensive and embarrassing (Young, 1991).

Performance Anxieties

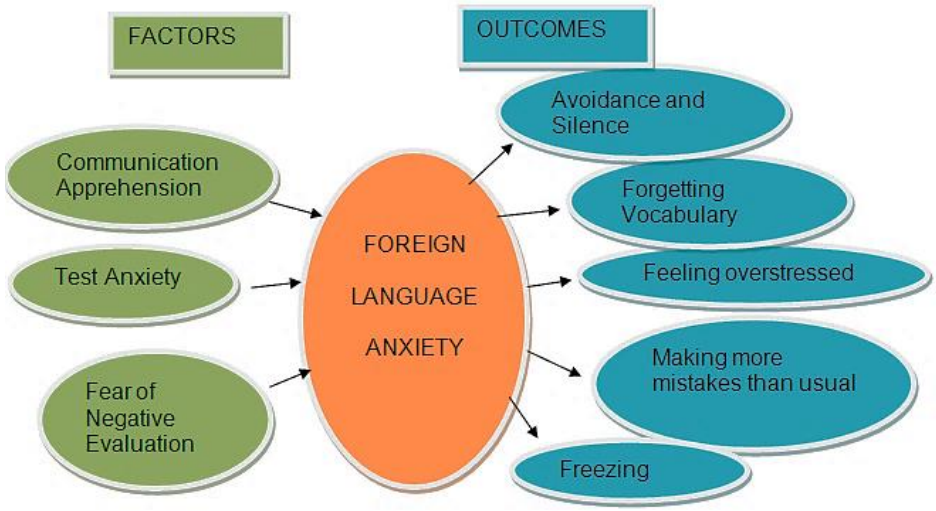
While some studies considered performance anxieties, i.e., communication apprehension, test anxiety, and fear of negative evaluation, as different types of language anxiety, other studies viewed them as components of language anxiety (Aydin et al., 2006; Tanveer, 2007). This study considers performance anxieties as factors contributing to language anxiety (See Figure 2).

Communication Apprehension

McCroskey and colleagues (1985) define communication apprehension as the anxiety or fear an individual experiences when communicating with others, either in real-life or anticipated situations. Language learners experienced communication apprehension when they believed their language proficiency was insufficient to convey their ideas clearly (de Blakeley et al., 2017; Horwitz et al., 1991; Thompson & Lee, 2013). This lack of confidence caused learners to refrain from speaking in language classes and shy away from conversing with more proficient speakers (Conway, 2007; McCroskey et al., 1985; Pappamihiel, 2002; Tsang, 2020). For example, Mustapha, Ismail, and Singh's (2010) study found that 45% of students studying Business Administration at a Malaysian college avoided public speaking and felt tense and nervous when communicating in group discussions due to high communication apprehension. As a result, communication apprehension may cause learners to avoid speaking in the classroom and experience high anxiety when asked to communicate with others (Lee, 2009; Morita, 2004).

Figure 2

Performance Anxieties as Factors Provoking Language Anxiety and Some Possible Outcomes



Test Anxiety

Test anxiety is a form of performance anxiety that can contribute significantly to foreign language anxiety (Cheng et al., 2014). It is common among language learners who believe that their language skills are being tested, and they may fail in this test. Test anxiety can cause language learners to experience high levels of stress before an exam, leading to forgetfulness and difficulty recalling vocabulary and grammar during the test (Horwitz et al., 1991). This can also result in language learners feeling overstressed and frustrated and even crying during the tests due to fear of failure (MacIntyre & Gardner, 1994). Because of test anxiety, language learners may perform poorly on language tests, exacerbating their future language anxiety (Horwitz et al., 1986). This is particularly true when language learning is compulsory, and success on language tests is necessary to enroll in college or obtain a job (Zheng, 2008). Test anxiety is, therefore, a significant contributor to foreign language anxiety among language learners.

Fear of Negative Evaluation

The fear of negative evaluation is a type of performance anxiety that arises from the fear of being judged or making mistakes (Giray et al., 2022; Hanh, 2020; Hamouda, 2013). This type of anxiety is closely linked to communication apprehension and can be especially detrimental to language learners. Language learners may feel ashamed of their language skills and worry about making grammar or pronunciation mistakes, which can increase their language anxiety (Aydin, 2008; Mesri, 2012; Ohata, 2005). Studies have demonstrated that the fear of negative evaluation is a major contributor to foreign language anxiety (Aydin, 2008; Kitano, 2001; Koralp, 2005), and it can impact communication with teachers and classmates and attitudes toward language learning. It is essential to consider and address the fear of negative evaluation to address foreign language anxiety among language learners.

The fear of negative evaluation pertains to the apprehension caused by perceived disapproval or criticism from others for not meeting their expectations. Individuals may experience fear of negative evaluation when they sense negativity or disapproval from others (Giray et al., 2022). This fear can be particularly detrimental for language learners, as they may feel that those around them are evaluating their language skills, competence, and performance. Consequently, language learners may become overly concerned about making a good impression on others and fear that they will be humiliated or criticized if they make mistakes (Capan & Simsek, 2012; Mesri, 2012; Ohata, 2005).

Several studies have investigated the impact of the fear of negative evaluation on language learners' performance in the classroom (Tatar, 2005). For example, Capan and Simsek (2012) found that many Turkish EFL students reported feeling afraid to speak in class due to fear of making mistakes in front of others. Additionally, Price (1991) and Doyman and Yumru (2020) found that anxious students who compared their language skills to their peers and felt lower proficiencies were likelier to feel anxious and experience low classroom participation. These studies highlight how much the fear of negative evaluation can affect language learners' classroom participation and performance. Therefore, researchers and educators need to pay more attention to this factor in decreasing language learners' anxiety levels.

Factors Contributing to the Fear of Negative Evaluation

The factors contributing to the fear of negative evaluation among language learners include language learners' personality traits, teachers' language ideologies, classroom dynamics, perfectionist tendencies, learners' multilingualism, and learners' ages and language proficiencies (de Blakeley et al., 2017; Dewaele, 2007; Dewaele et al., 2008; Jee, 2022; Oad et al., 2020; Thompson & Khawaja, 2016). These factors can be categorized under two themes: individual factors and classroom-related factors.

Individual Factors

Perfectionist tendencies of language learners can be viewed as an individual factor contributing to the fear of negative evaluation among language learners. These tendencies may lead learners to feel threatened by the negative judgments of others (Gregersen & Horwitz, 2002; MacIntyre & Gardner, 1991). For example, Xianping (2003) found that perfectionist language learners were seldom satisfied with their achievements and were far more concerned about their limitations than non-perfectionist language learners. Zheng (2008) added that language learners with perfectionist tendencies believed they always had to speak native-like to avoid being scrutinized. After interviewing eight college-level L2 learners, Gregersen and Horwitz (2002) suggested that perfectionist

language learners' anxiety levels spiked substantially when speaking in front of others in the target language since it was considered a risk-taking activity. Thus, perfectionist tendencies may increase L2 learners' fear of negative evaluation (Horwitz et al., 1986).

Lack of preparedness for classroom discussions can also contribute to language learners' fear of negative evaluation. Horwitz, Horwitz and Cope (1986) observed that students tended to become tongue-tied in situations that demanded spontaneous language use but found comfort in rehearsed speeches and drills. Students may feel anxious about being called on because they fear making mistakes and receiving negative evaluations from teachers or peers (Pappamihiel, 2002). Therefore, language learners may experience high levels of fear of negative evaluation due to their lack of preparedness for spontaneous language production (Fandiño Parra, 2010).

Self-perceived language ability or language ability mindsets can serve as a gauge for fear of negative evaluation among language learners (Kitano, 2001; Saito & Samimy, 1996; Wu et al., 2015). Young (1991) and Jee (2022) contend that students who perceived their language proficiency as low were more vulnerable to fear of negative evaluations in class. Ozdemir and Papi (2022) found that students' beliefs about the malleability of their language learning intelligence influenced their anxiety and fear of negative evaluation. Kitano's (2001) and Zarei and Moussavou's (2022) studies revealed that the second language learners' levels of fear of negative evaluation relate to their self-perceived speaking skills and that this correlation also impacted their language anxieties. Thus, language learners who deemed their communicative competence in the target language inadequate to express themselves or communicate with the native speakers were likely to experience high levels of fear of negative evaluation in the classroom.

Age could also be considered a factor affecting the fear of negative evaluation (de Blakeley et al., 2017; Santos et al., 2017). Frost et al. (1995) revealed that low self-esteem led to a higher fear of negative evaluation, and Orth et al. (2010) argued that people generally gain self-esteem as they grow older. Aydin (2008) noted that younger language learners tended to worry about their peers' opinions and were often preoccupied with pleasing others in foreign language classes compared to older students (Young, 1992). Consequently, higher self-esteem in older adults may represent a lesser level of fear of negative evaluation (Santos et al., 2017).

Finally, gender may play a role in determining language learners' fear of negative evaluation (Dewaele, 2007). However, the role of gender in fear of negative evaluation is controversial. While some studies revealed a lower threshold of fear of negative evaluation among females (Mesri, 2012; Ozturk & Gurbuz, 2013; Ran et al., 2022), others suggested that males suffer from the fear of negative evaluation more than females (Capan & Simsek, 2012; Zheng, 2008). Pappamihiel (2002) argued that males are less likely to acknowledge their fears, which could be one reason why some studies have found females to be more afraid of being judged than males. As a result, males and females might differ in experiencing fear of negative evaluation in language classes.

Classroom-Related Factors

The fear of negative evaluation can also be increased due to the teachers' practices (Horwitz et al., 1986; Pappamihiel, 2002; Yan & Horwitz, 2008). Teachers' constructive feedback or error corrections can give language learners the impression that they are incapable of mastering the target language. Von Wörde (2003) discovered in his study involving 15 language learners that students experienced irritation and frustration, particularly when teachers started to admonish them for their errors. Additionally, Capan and Simsek's (2012) research, which involved 131 Turkish college-level English language learners, found that many students feared making mistakes due to teachers' potential negative attitudes or feedback. Therefore, teachers' practices may increase the fear of negative evaluation if students perceive any threats to their faces.

The presence of native speakers or more proficient peers in a shared classroom may provoke the fear of negative evaluation among language learners (Baker & MacIntyre, 2000; Conway, 2007; Doyman & Yumru, 2020; Ohata, 2005; Pappamihiel, 2002). For instance, many language learners expressed concerns about their performance during classroom activities and oral presentations due to the fear of negative judgments or humiliation from more proficient peers in Ohata's (2005) and Young's (1991) studies. Similarly, Conway (2007) claims that "anxious students compare their skills with native speakers of the target language, which leads to embarrassment and shame when they are not pronouncing exactly like the native speakers" (p. 5). Similarly, Pappamihiel (2002) reported higher levels of fear of negative evaluation among ESL students in mainstream classrooms rather than in ESOL classrooms due to the presence of native speakers in mainstream classrooms.

Some teacher practices can decrease students' fear of negative evaluation and anxiety. For example, Patra et al. (2022) revealed in their study with 76 male EFL learners that corrective feedback diminished anxiety and increased students' engagement. Fattahi-Marnani and Cuocci (2022) suggested that teachers know their students' ethnic, cultural and linguistic backgrounds better to relate the topics to their lives. Seltzer and de los Ros (2018) suggested that teachers design critical discussions in their study so that students from different linguistic and cultural backgrounds could connect with each other. Finally, Russell (2020) suggested that teachers can provide EFL students with learning tips, have them do some relaxation activities such as deep breathing, assigned tasks such as role plays or journaling to express their emotions in speaking a foreign language and construct support teams.

The Impacts of Fear of Negative Evaluation on Student Engagement

The fear of negative evaluation can reduce student engagement as it can cause students to become overly worried about making mistakes and being judged by others (Morita, 2004; Tatar, 2005). Specifically, speaking activities may be perceived as threatening for language learners who experience this fear, leading to a reluctance to initiate conversations and minimal interaction during group discussions (Xianping, 2003; Zarei & Moussavou, 2022). Studies showed that students with high levels of fear of negative evaluation often remained passive in class, avoided participating in activities that could help them practice their speaking skills, and allowed them opportunities to demonstrate their knowledge (Xianping, 2003).

The fear of negative evaluation can extend its adverse impact on language output. Individuals harboring this fear often display an intensified focus on the precision of language production. Paradoxically, their excessive concern about avoiding mistakes may lead to an elevated occurrence of errors in their linguistic output (Kasap, 2023; Ludwig, 1982; Pak, 2014). Xianping's (2003) research, which involved interviews and observations of 97 non-English major college students, revealed that those with high levels of fear of evaluation tended to speak less, produce shorter sentences, and make more errors in their output. Similarly, Ran et al. (2022) noted that students with high levels of fear of negative evaluation spoke slowly, frequently forgot vocabulary and grammar rules, and could not convey their messages in the target language. Finally, Nikolov and Djigunovic (2006) revealed that students had "long mid-clause pauses, and rarely repeated the main points and started with false discourse to the conversation" due to the high levels of fear of negative evaluation (p. 246).

Methodology

Research Design

This qualitative research study is designed as a multiple case study. Multiple case studies help researchers accumulate a rich and diverse set of data, enhancing the depth of understanding of the subject under investigation (Crowe et al., 2011). Moreover, this design gives researchers a more holistic understanding of a complex phenomenon. This multi-case study aimed to explore the international graduate students' languaging and learning experiences in classrooms involving native English speakers (NES) peers. Specifically, the study explored the role of native-speaker peers in students' fear of negative evaluation and classroom participation. Thus, it sought answers to the following questions:

1. How does the presence of native-speaker peers influence international graduate students' fear of negative evaluation?

2. How does the presence of native-speaker peers influence international graduate students' participation?

The research methodology employed in this study is a component of my master's thesis submitted to the University of Florida. It is important to note that while this manuscript focuses on specific aspects of the larger thesis, the methodology outlined here represents a key element of the comprehensive research conducted for the completion of my master's degree.

Context

This study took place in a public university in the U.S.A. International students who studied at this university had to score above 90 on the TOEFL IBT test to get admitted to the graduate programs in this university. Moreover, to get accepted to graduate programs at this university, all students had to score above a certain score on the GRE test, which measures students' reading, writing, and analytical thinking skills. Graduate programs required

students to take compulsory and elective courses. While compulsory courses were offered only to students in specific programs, elective courses were offered to students studying in different disciplines. Based on students' interests and familiarity with the content, while some courses were chosen mainly by native English students (e.g., history of American Education), others were chosen by mostly international students (e.g., Cross-cultural Communications).

Participants

The survey participants were selected with convenience sampling as the first author distributed the survey link to people in her social network. This sampling strategy implied that everyone in the population did not have an equal chance of being selected, and thus, the findings of the survey results cannot be generalized to the whole NNS graduate student population in the USA. As the function of the survey was to determine interview participants and not propose any generalizable hypothesis, convenience sampling was considered appropriate for this study. The survey was given to 22 graduate students studying at various universities in the United States. The participants were from diverse ethnic backgrounds, including Chinese (9), Turkish (5), Arab (2), Korean (2), Persian (1), Japanese (1), Dutch (1), and German (1). Sixteen of the participants were female, and six were male. The age range of participants varied from 18 to 55, with 59% falling within the 26-35 age group. A significant majority (82%) of participants were pursuing graduate studies in ESOL/Bilingual Education or Linguistics. All participants had initially learned English as a foreign language in their home countries and currently spoke it at an advanced level as a second language. Before recruitment, it was confirmed that all participants took courses in classes where the majority of students were native speakers and classes where the majority were non-native speakers.

Interview participants were selected using a purposive sampling method (Palinkas et al., 2015). From the survey participants, five participants exhibiting a notable fear of negative evaluation in classrooms predominantly composed of native speakers were selected for individual interviews to delve deeper into their experiences in such settings. All participants were female, with three—Bao, Yin, and Zhen—being of Chinese descent, Alina being Arab, and Daria being German. Four of the participants—Bao, Yin, Zhen, and Daria—were graduates. Students in the ESOL/Bilingual Education program, while Alina pursued her graduate studies in the Linguistics department. All participants were fluent in English.

Data Collection and Analysis

The data collection tools of this study included only interviews. However, the researchers used an online survey only to determine the interview participants. The online survey was prepared in the Qualtrics program. The survey included 60 items and was prepared as a combination and adaptation of the Foreign Language Classroom Anxiety Scale (FLCAS) designed by Horwitz (1986), the English Language Anxiety Scale (ELAS) (Pappamihiel, 2002), the Japanese Class Anxiety Scale developed by Kitano (2001), Fear of Negative Evaluation Scale (Leary, 1983) and the researcher's personal experiences. Since this study specifically examined the role of native speakers on fear of negative evaluation and there was no scale that served this purpose, different items that focused on native speakers' presence were selected from the abovementioned scales. The survey data were analyzed with descriptive statistics. The five students who scored higher mean scores in certain items ($m > 3.7$ out of 5.0) were selected for the interviews as their scores implied a high level of language anxiety in classrooms with predominantly native English speakers.

Semi-structured interviews were conducted with each participant to explore international students' experiences in classrooms with predominantly NES students. Each interview with the five participants involved 10 to 12 open-ended questions and took 35-45 minutes. The questions included items regarding NES students' attitudes to their presence or contributions, how these attitudes influenced their academic or linguistic performance and instructors' inclusive practices and interactions with them. These items were determined based on participants' answers to the survey questions and the purpose of this study. The interviews were audio-recorded to be transcribed and analyzed later with the "Domain Analysis" technique (Spradley, 1979).

Domain analysis was selected because we considered graduate classrooms as one domain that had different demographics (involved more culturally and linguistically diverse students compared to undergraduate programs) and jargon as they included more academic and critical discussions. The first step of the data analysis process involved identifying the elements related to the fear of negative evaluation in the data and coding them with key terms such as biased attitudes, language proficiency as power and misinterpretations of international students' contributions. The second step included organizing the key terms into categories based on the elements

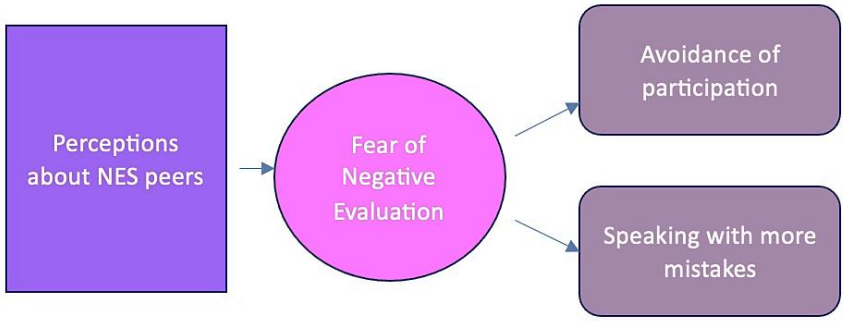
of graduate programs, such as instructors, peers, and pedagogies. Finally, the related key terms, such as perceptions, participation, and production, were compiled under the cover terms and presented as themes.

Findings

The findings revealed that the presence of NES peers influenced international graduate students' fear of negative evaluation and participation in class activities. The data showed that international students' perceptions of their NES peers' attitudes and practices determined their fear of negative evaluation. Based on their perceptions, this fear influenced both their engagement with the activities and languaging practices (see Figure 3).

Figure 3

The Impacts of NES Presence on Fear of Negative Evaluation



International Students' Perceptions of the NES Peers' Attitudes and Practices

All participants highlighted that their NES peers were generally benevolent towards them and did not judge them based on their language proficiency. However, they also agreed that some NES peers demonstrated biased attitudes toward international students, often increasing their fear of negative evaluation. Zhen explained the negative attitudes of her NES peers as such:

> Students in the elementary education program, some of them are really bad… in terms of their attitude to international students and they, I guess, intentionally give us the pressure, give us the fear like they don't care about us, and they don't want to hear our ideas or our decisions. They don't want to interact with us in group discussions. They just express what they think as group idea. That makes me feel really bad, so I don't like the classes I had with students in the elementary education program (96-98).

Daria experienced similar discriminatory attitudes from some of her NES peers, and claimed that international students' non-native accents were the main reason for these attitudes. She said:

> With native speakers, I am afraid that they don't understand me first of all because of my accent, because I experienced that people asked me 'what?', and then I repeated, and they said 'what?' again. And that happened maybe three times, so that happened often. And in that cases, I try to speak more pronounced, but at times people couldn't understand me, I don't know why. So that was pretty uncomfortable situation, and people in this academic language, people are speaking on the very high level, and I just try to, I just feel I have to match that somehow (151).

Alina claimed that some of her NES peers positioned her as cognitively incompetent due to her linguistic 'deficiencies.' She said:

> I feel like when they are listening to me, if you are some non-native or your sentence is full of mistakes, I feel like they perceive you as a person who is, I don't want to say stupid, but you need to speak more English, or you need to learn before speaking (8).

These negative attitudes often increase ESL students' fear of negative evaluation. Zhen explained the situation as such:

> If they (NES peers) try to help and accept us as who we are and how our English is, we will have lower anxiety. If you can sense from her or his facial expression, 'oh maybe he or she doesn't like my English and doesn't want to pay attention', we say, 'oh my God, I have to try really better to express my ideas'. So when I have teammates who are native English speakers, they didn't pay attention to what I am saying, I feel a little bit pissed off, frustrated, and also nervous … They don't care what international students think (90).

In sum, international students perceived negative vibes from some of their NES peers due to their accents or identities. When their attitudes or practices threatened their faces or positioned them cognitively inferior, their fear of negative evaluation increased.

The Impact of the NES Presence on International Students' Participation

Due to their linguistic advantage, the international students agreed that their NES peers were one step ahead of them in learning or participation. Daria exemplified the situation as such:

> Let's say some philosophical discussions about something, they, for sure, are advantaged, and I mean before I have formed my thoughts about something on the high level like this, at least myself, ... you have them right there. Just being a native speaker is a big advantage; you have two pieces together (181).

Yin agreed with Daria and noted that "sometimes, international students, non-native students, they will also have very quick ideas, but due to the language barrier, they may have a hard time explaining their ideas" (323).

Instructors' interactions with NES peers also influenced ESL students' engagement with the content. All participants agreed that instructors interacted with their NES peers more often than the international students. Alina shared her observation about instructors' attitudes as such: "Professors sometimes feel more comfortable talking with the native speakers because they get the answer quickly, and they can understand any small details when they are saying... They are looking at them more than they are looking at non-native speakers" (14). Similarly, Daria complained that instructors provided more compliments to her NES peers than she did, although she tried very hard to be successful in the courses.

Instructors' limited comprehension or ignoration of international students' ideas also influenced their fear of negative evaluation. Daria and Zhen claimed that instructors had difficulties understanding the international students' points due to their linguistic 'limitations' and rarely asked for clarification, implying that they did not value them. Zhen said:

> One thing I found always is that they didn't understand what we were talking about. Most of the time, they won't ask questions to clarify ideas. So they will just rephrase what we said, but that's not exactly our words or our thoughts. That's just how they think we said. So we had different ideas. What the professor said was not our words. They did not understand, but they didn't want to clarify, I don't know why. Maybe they don't want to make us feel bad, or they just don't want to waste time listening to us (96).

The discussion topics that required a high level of academic language were another point that influenced ESL students' engagement and fear of negative evaluation. Daria expressed that the philosophical topics requiring a high level of academic language put her under stress because she felt that her academic English language skills sometimes remained insufficient to understand and discuss these topics (177). Similarly, Alina said she would avoid discussing topics such as religion that made her uncomfortable because they could leave an unfavorable impression on her NES peers. Thus, discussion topics discussed in the classrooms could also influence ESL students' fear of negative evaluation.

All participants noted that the presence of NES peers influenced their engagement and willingness to participate considerably because they feared being judged by them when they made a mistake. For example, Daria proposed that she wanted to speak in a native-like manner, especially with her pronunciation, to leave good impressions on her NES peers or professors (157). Thus, they claimed they would participate only when they were confident in their languaging practices. For example, Alina emphasized that she spoke whenever she was sure that her sentences were accurately formed because she did not want to sound stupid. Zhen said:

> What I observe is all classmates, when we had more native speakers, we would prefer not to participate. I don't know whether because we don't know the answer, or like me, we are just afraid of making mistakes (88).

Finally, they expressed that they would participate in lessons more if they did not fear making mistakes. For example, Yin said: "If I didn't really care how others judge my pronunciation or my proficiency, I would like volunteering more, I guess" (345).

In conclusion, all participants acknowledged that the presence of NES peers in their classrooms influenced instructors' practices, linguistic requirements of the discussion topics, and their feelings about fear of making mistakes and being judged. International students perceived the privilege of NES peers' practices in classrooms with NES peers and felt more fear of negative evaluation as they practiced differently from them. These varying practices influenced their participation and engagement in the lessons.

The Impact of the NES Presence on International Students' Languaging Practices

Most participants felt comfortable interacting with other international and NES people in their social network. For example, Alina said, "When I am speaking with my friends, I feel more free. I can use more idioms. I feel like I am more like a native" (8). Zhen explained why she felt less anxious around other international students: "I feel less anxious. I don't know why, I guess because I think we are all learners, and it's normal to make mistakes."

On the other hand, they felt very anxious and worried about being negatively judged when interacting with their more proficient international and native English-speaker classmates. These feelings, in return, influenced their languaging practices negatively and prevented them from demonstrating their full linguistic potential. For example, Daria said:

> I feel very much that I have to speak very well and also my pronunciation on one hand, I feel like I try to suppress my accent, my German accent, but then when I try harder, it comes out more. Because I am more tense, and I am less… I think I speak better English if I'm more relaxed. Then, the American pronunciation would come out better. Then being tense causes me to pronounce it more hard like a German word. (157).

She added the following to explain how her sentences were influenced by her anxiety as such:

> It's the vocabulary…, I cannot find the right word in English. And then, I am afraid that my sentences are kind of very simple maybe, you know, because if you try to explain something difficult or use simple words, it may sound like, you sound like low-level English, you know… when I then use simple words to explain something, that's when I feel all that may be sounding very bad, or I mix the forms, maybe I use the present form, or I make a mistake maybe with the tenses or sentence structure sometimes depending (193).

To sum up, international students had less anxiety around other international students than their NES peers. Moreover, international students' fear of negative evaluation influenced their languaging practices negatively.

Discussion

The findings revealed that the international students' fear of negative evaluation increased, whereas their willingness to participate in whole-class discussions decreased in the presence of native English speaker peers in the classrooms. The fear of negative evaluation among international learners was mainly influenced by their perceptions of their native English-speaking (NES) students' attitudes and practices. The international graduate students often harbored a pervasive belief that their NES counterparts perceived them as outsiders and as academically less proficient, largely on the basis of their discernible international accents (Ozdemir & Papi, 2022; Tan et al., 2021). This perception led to a pronounced sense of otherization within the classroom environment (Coppinger & Sheridan, 2022). Thus, in settings predominantly populated by NES peers, international students commonly experienced heightened levels of anxiety. This dynamic underscores the importance of creating inclusive, safe, supportive educational environments in graduate schools that foster cultural understanding and appreciation for linguistic diversity (Bruen & Kelly, 2017; Tercan & Dikilitas, 2015). Addressing these perceptions and fostering an atmosphere of acceptance and mutual respect is imperative, ultimately enhancing the overall learning experience for all students involved (Ozdemir & Papi, 2022).

Instructors' practices were the second factor that influenced the international students' fear of negative evaluation, as also found in Malik et al. (2020). Participants in the study noted that certain aspects of classroom practices, such as the way teachers corrected errors, their interactions with international and NES students, and the topics covered in group and class discussions, could negatively impact international students' fear of negative evaluation. Horwitz, Horwitz, and Cope (1986) found that specific teacher practices, such as failing to

acknowledge the achievements of international students, engaging in warmer social interactions with NES students, or displaying varied attitudes during corrections, can increase international students' anxiety and influence their language production, and class participation. Moreover, the unfamiliarity of international students with the discussed topics can influence their language use and occasionally decrease engagement. These findings implied that academic staff had an important role in international students' attitudes to courses and engagement with the content. Thus, academic staff who had international students in their classrooms need to design their curriculum, teaching practices and interaction patterns in an inclusive way that linguistic diversity is celebrated and mistakes are seen as valuable learning opportunities rather than sources of anxiety.

The fear of negative evaluation impacted the classroom participation of international students considerably, primarily due to their fear of being judged by their NES peers who had linguistic power in the classrooms. This erratic power dynamic often created a perceived discrepancy in communication proficiency, leading international students to exhibit reticence in participating actively in classroom discussions, as also found in Ran et al. (2022). They may have felt a sense of apprehension or inadequacy and insecurity in their linguistic or academic abilities around NES students as they positioned NES students more knowledgeable (Doyman & Yumru, 2020; Russell, 2020; Tan et al., 2021). When instructors' practices, such as more positive feedback to NES students' contributions, confirmed this belief, their fear of negative evaluation was increased, and international students behaved abstentionist, as also revealed in Ran et al.'s (2022) and Fattahi-Marnani and Cuocci's (2022) studies. As a result, it is essential for educators to implement strategies that level the playing field, ensuring that all students, regardless of their linguistic background, have equal opportunities to contribute meaningfully to the learning environment. Moreover, instructors can engage students in critical discussions or create support teams that build connections between NES and international students and reveal international students' expertise, as suggested by Seltzer and de los Rios (2018) and Russell (2020).

Lastly, the presence of NES peers exerted a discernible influence on the languaging practices of international students, with notable variations in response (Ludwig, 1982). When in the company of NES peers, some students experienced a heightened fear of errors, driven by the desire to attain a level of linguistic proficiency akin to that of a native speaker (Pak, 2014). This pursuit of perfection inadvertently exacerbated their anxiety levels, leading to an increase in linguistic missteps (Ran et al., 2022; Xianping, 2003). Over time, this dynamic began to take a toll on their physical well-being within the classroom environment. These findings underscore the complex interplay between language anxiety, peer dynamics, and the resultant impact on students' overall well-being. The findings implied a need for critical discussions that center language learners' feelings in NES-dominant classrooms so that NES students can be more inviting, respectful, and less judgmental in their interactions with international students. Russell (2020) suggested support teams in classrooms where students with higher levels of fear of negative evaluation can learn from peers and get encouraged in their language productions. A study group composed of NES students can also develop relations between international and NES students.

Conclusion

The findings of this research indicate a prevalent occurrence of the fear of negative evaluation among graduate-level international learners. This study sought to elucidate the experiences of international students who exhibited heightened levels of this fear in response to the presence of native speakers in the classroom. It is noteworthy to acknowledge that native speakers themselves may also experience a fear of negative evaluation, potentially stemming from various factors. This apprehension could impede their active participation and verbal communication. Given this phenomenon's complexity and substantial impact on students' inclination to engage in classroom activities and oral expression, further investigation is warranted.

The findings suggest that the apprehension of international graduate students towards negative evaluation was shaped by their interactions with native speakers, as well as the perfectionist tendencies of either native speakers or international learners, along with classroom methodologies, as also found by Ozdemir and Papi (2022). Additionally, the study demonstrated that, apart from the fear of negative evaluation, personal traits like personality, perfectionism, and readiness had an adverse impact on L2 learners' classroom engagement. Lastly, it was proposed that a heightened fear of negative evaluation impeded the accuracy of oral language expression.

The findings imply that creating a safe environment where linguistic and cultural diversity are valued, international students' home languages can be used, and linguistic mistakes are perceived as opportunities to learn is important in graduate schools that involve international students (Aptoula, 2022). Academic staff need to adopt more inclusive approaches in their interactions with their NES and international students and determine the

discussion topics (Fattahi-Marnani & Cuocci, 2022). They can also provide international students with corrective and positive feedback or language learning tips to decrease their anxiety (Giray et al., 2022; Patra et al., 2022). Finally, since NES peers' attitudes had an essential role in provoking international students' fear of negative evaluation., it is crucial to open discussions in the classrooms in which international students' feelings can be negotiated, and NES students' awareness can be increased (Seltzer & de los Rios, 2018). Creating a support group composed of NES students can also build connections among NES and international students and decrease fear of negative evaluation (Russell, 2020).

To sum up, this research enriches the existing literature by shedding light on the experiences of international learners in classrooms predominantly occupied by native English-speaking (NES) students. The study underscores that the presence of NES students may trigger a heightened fear of negative evaluation among international learners, leading to more adverse consequences compared to classrooms primarily comprised of international students. Considering these findings, educators in such settings need to be cognizant of the potential negative impacts of NES presence and take proactive measures to mitigate the influence of fear of negative evaluation arising from their presence.

Future Research

Although this study revealed significant findings regarding the learning environments and peer interactions in classrooms, predominantly NES students, it had limitations. Future research studies can involve more students with high levels of anxiety, not only at the graduate level but also at the undergraduate level. Moreover, involving international students with low-level anxiety can be an essential contribution as it can highlight exemplary practices that comfort international students. Lastly, studies that involved international students in other countries where English is spoken as the international language can reveal essential findings to understand dynamics in the classrooms where the power relations differ from graduate schools in the U.S.A.

This research specifically addresses the adverse impact of native speaker presence on classroom involvement and the proficiency of language expression, relying solely on individual interviews. The information gleaned from these interviews solely represents the viewpoints and conjectures of the participants. To address this limitation, subsequent studies could explore the same issue through additional data collection methods. This approach would offer researchers a variety of vantage points to analyze the situation and enable a comparison of students' perspectives with their actual experiences. It also opens the door to incorporating the perspective of educators, potentially yielding more robust and reliable findings.

Finally, further investigation into the fear of negative evaluation arising from the presence of native speakers is crucial to supporting international learners in ESL environments. This phenomenon has also been indicated by numerous studies, including the current one, to affect the classroom behaviors of international learners adversely (Aydin, 2008; Baker & MacIntyre, 2000; Horwitz et al., 1986; Kitano, 2001; Pappamihiel, 2002). Additionally, various elements such as teachers' attitudes, learners' perfectionist tendencies, personality traits, preparedness, and the attitudes of native speakers can impact the levels of fear of negative evaluation experienced by international learners (Shabani, 2012; Kitano, 2001; Pappamihiel, 2002). Researching any of these factors further could substantially contribute to the existing body of literature.

Code of Ethics

This research study was approved to be ethically appropriate with the decision of the Behavioral/NonMedical Institutional Review Board (IRB) of the University of Florida, dated on August 28, 2013, and numbered IRB201700418.

References

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's constructs of foreign language anxiety: The case of students of Japanese. *Modern Language Journal, 78*, 155–168. https://doi.org/10.2307/329005

Aptoula, N. Y. (2022). Exploring academic literacy practices of graduate students in English language teacher education programmes at English-medium universities in Turkey. *Literacy, 56*(2), 174–183. https://doi.org/10.1111/lit.12279

Aydin, S. (2008). An investigation on language anxiety and fear of negative evaluation among Turkish EFL learners. *Asian EFL Journal, Teaching Articles, 30(*1), 421–444. https://files.eric.ed.gov/fulltext/ED512266.pdf

Aydin, S., Yavuz, F., & Yesilyurt, S. (2006). Test anxiety in foreign language learning. *Journal of Graduate School of Social Sciences of Balikesir University, 9*(1), 145–160. https://files.eric.ed.gov/fulltext/ED511092.pdf

Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning, 50*(2), 311–341. https://doi.org/10.1111/0023-8333.00119

Bruen, J., & Kelly, N. (2017). Using a shared L1 to reduce cognitive overload and anxiety levels in the L2 classroom. *The Language Learning Journal, 45*(3), 368–381. https://doi.org/10.1080/09571736.2014.908405

Capan, S. A., & Simsek, H. (2012). General foreign language anxiety among EFL learners: A survey study. *Frontiers of Language and Teaching, 3*, 116–124.

Cheng, L., Klinger, D., Fox, J., Doe, C., Jin, Y., & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET and GEPT. *TESOL Quarterly, 48*(2), 300–330. https://doi.org/10.1002/tesq.105

Conway, J. (2007). *Anxiety in second language learning: Causes and solutions.* [Unpublished paper]. http://www.academia.edu/5425739/380Paper_Jennifer

Coppinger, L., & Sheridan, S. (2022). Accent anxiety: An exploration of non-native accent as a source of speaking anxiety among English as a foreign language (EFL) students. *Journal for the Psychology of Language Learning, 4*(2), 1–21. https://doi.org/10.52598/jpll/4/2/6

Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical Research Methodology, 11*, 100. https://doi.org/10.1186/1471-2288-11-100

de Blakeley, M. G., Ford, R., & Casey, L. (2017). Second language anxiety among Latino American immigrants in Australia. *International Journal of Bilingual Education and Bilingualism, 20*(7), 759–772, https://doi.org/10.1080/13670050.2015.1083533

Dewaele, J. M. (2007). The effect of multilingualism socio-biographical and situational factors on communicative anxiety and foreign language anxiety of mature language learners. *The International Journal of Bilingualism, 11*(4), 391–409. https://doi.org/10.1177/13670069070110040301

Dewaele, J. M., Petrides, K. V., & Furnham, A. (2008). The effects of trait emotional intelligence and socio-biographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning, 58*(4), 911–960. https://doi.org/10.1111/j.1467–9922.2008.00482.x

Doyman, S., & Yumru, H. (2020). An exploration of students' perceived sources of speaking anxiety. *International Journal of Media, Culture and Literature, 6*(2), 189–199. https://doi.org/10.17932/IAU.IJMCL.2015.014/ijmcl_v06i2005

Eysenck, M. W. (1979). Anxiety, learning, and memory: A reconceptualization. *Journal of Research in Personality, 13*(4), 363–385. https://doi.org/10.1016/0092-6566(79)90001-1

Fandiño Parra, Y. J. (2010). Explicit teaching of socio-affective language learning strategies to beginner EFL students. *Íkala, Revista De Lenguaje Y Cultura, 15*(1), 145–169. https://doi.org/10.17533/udea.ikala.5099

Fattahi Marnani, P., & Cuocci, S. (2022). Foreign language anxiety: A review on theories, causes, consequences and implications for educators. *Journal of English Learner Education, 14*(2), 1–29. https://stars.library.ucf.edu/jele/vol14/iss2/2

Frost R. O., Turcotte T. A., Heimberg R. G., Mattia J. I., Holt C. S., Hope D. A. (1995). Reactions to mistakes among subjects high and low in perfectionistic concern over mistakes. *Cognitive Therapy and Research, 19*, 195–205. https://doi.org/10.1007/BF02229694

Giray, L., Alcala, M. A., Edem, J., & Sabacajan, T. M. (2022). English language anxiety among college students. *International Journal of Qualitative Research, 2*(1), 65–76. https://doi.org/10.47540/ijqr.v2i1.569

Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal, 86*, 562–570. http://www.jstor.org/stable/1192725

Hamouda, A. (2013). An exploration of causes of Saudi students' reluctance to participate in the English language classroom. *International Journal of English Language Education, 1*(1), 17–34.

Hanh, N. T. (2020). Silence is gold?: A study on students' silence in EFL classrooms. *International Journal of Higher Education, 9*(4), 153–160. https://doi.org/10.5430/ijhe.v9n4p153

Hewitt, E., & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillips's MLJ study. *The Modern Language Journal, 96*(2), 170–189. https://doi.org/10.1111/j.1540-4781.2011.01174.x

Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly, 20*, 559–564. https://doi.org/10.2307/3586302

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*, 125–132. https://doi.org/10.2307/327317

Horwitz, M. B., Horwitz, E. K., & Cope, J. (1991). Foreign language classroom anxiety. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 27–39). Prentice Hall.

Jee, M. J. (2022) Heritage language anxiety and major language anxiety experienced by Korean immigrants in Australia, International *Journal of Bilingual Education and Bilingualism, 25*(5), 1713–1729, https://doi.org/10.1080/13670050.2020.1799321

Kasap, S. (2023). The impact of musical skills on foreign language anxiety. *European Journal of English Language Teaching, 8*(2), 1–16. https://oapub.org/edu/index.php/ejel/article/view/4737/7372

Kasap, S., & Power, K. M. (2019). Anxiety in the EFL speaking classrooms. *The Journal of Language Teaching and Learning, 9*(2), 23–36.

Kitano, K. (2001). Anxiety in the college Japanese language classroom. *The Modern Language Journal, 85*(4), 549–566. https://doi.org/10.1111/0026-7902.00125

Koralp, K. (2005). *A retrospective analysis of the English language learning anxiety experienced by prospective teacher of English* [Unpublished doctoral dissertation]. Eastern Mediterranean University.

Krashen, S. (1998). Comprehensible output? *System, 26*, 175–182.

Leary, M. R. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin, 9*, 371–375. https://doi.org/10.1177/014616728309300

Lee, J. (2009). Self-constructs and anxiety across cultures. *ETS Research Report Series, 1*, i–35. https://doi.org/10.1002/j.2333-8504.2009.tb02169.x

Ludwig, J. (1982). Native-speaker judgments of second-language learners' efforts at communication: A review. *The Modern Language Journal, 66*(3), 274–283. https://www.jstor.org/stable/326629

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second language learning: Toward a theoretical clarification. *Language Learning, 39*(2), 251–275.

MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language Learning, 41*(4), 513–534.

MacIntyre, P., & Gardner, R. C. (1994). The Stable effects of language anxiety on cognitive processing in the second language. *Language Learning,* 44 (2), 283–305.

Malik, S., Qin, H., Khan, S. M., & Ahmed, K. (2020). Quantitative analysis of the foreign language anxiety: Chinese and Pakistani postgraduates in focus. *Arab World English Journal (AWEJ), 11*(1), 315–330. https://doi.org/10.2139/ssrn.3582338

McCroskey, J. C., Fayer, J. M., & Richmond, V. P. (1985). Don't speak to me in English: Communication apprehension in Puerto Rico. *Communication Quarterly, 33*, 185–192.

Mesri, F. (2012). The relationship between gender and Iranian EFL learners' foreign language classroom anxiety (FLCA). *International Journal of Academic Research in Business and Social Sciences, 2*(6), 147–156.

Morita, N. (2004). Negotiating participation and identity in second language academic communities. *TESOL Quarterly, 38(*4), 573–603. https://doi.org/10.2307/3588281

Mustapha, W. Z. W., Ismail, N., & Singh, D. S. R. (2010). ESL students' communication apprehension and their choice of communicative activities. *AJTLHE, 2*(1), 22–29

Nikolov, M., & Djigunović, J. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics, 26,* 234–260. https://doi.org/10.1017/S0267190506000122

Oad, L., Khan, N., Khoso, F. (2020). Factors affecting English language anxiety among learners: A case study of a public sector university. *Pakistan Social Sciences, 4*(3), 1059–1078. http://doi.org/10.35484/pssr.2020(4-III)77

Ohata, K. (2005). Potential sources of anxiety for Japanese learners of English: Preliminary case interviews with five Japanese college students in the U.S. *TESL-EJ, 9*(3), 1–21. https://eric.ed.gov/?id=EJ1065859

Orth, U., Trzesniewski, K. H., & Robins, R. W. (2010). Self-esteem development from young adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 98*(4), 645–658. https://www.apa.org/pubs/journals/releases/psp-98-4-645.pdf

Oxford, R. L. (1999). Anxiety and the language learner: new insights. In J. Arnold (Ed.). *Affect in language learning* (pp. 58–67). Cambridge University Press.

Ozdemir, E., & Papi, M. (2022) Mindsets as sources of L2 speaking anxiety and self-confidence: The case of international teaching assistants in the U.S. *Innovation in Language Learning and Teaching, 16*(3), 234–248, http://doi.org/10.1080/17501229.2021.1907750

Ozturk, G., & Gurbuz, N. (2013). The impact of gender on foreign language speaking anxiety and motivation. *Procedia-Social and Behavioral Sciences, 70*, 654–665.

Pak, C. (2014). *Speak up or be silent? Language learners' anxiety and motivation on speaking up in ELL and non-ELL classrooms* [Unpublished doctoral dissertation]. University of Washington.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health, 42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Pappamihiel, N. E. (2002). English as a second language students and English language anxiety: Issues in the mainstream classroom. *Research in the Teaching of English, 36*, 327–355. https://www.jstor.org/stable/40171530

Patra, I., Alazemi, A., Al-Jamal, D., & Gheisari, A. (2022). The effectiveness of teachers' written and verbal corrective feedback (CF) during formative assessment (FA) on male language learners' academic anxiety (AA), academic performance (AP), and attitude toward learning (ATL). *Language Testing in Asia, 12*, 19. https://doi.org/10.1186/s40468-022-00169-2

Price, M. L. (1991). The subjective experience of foreign language anxiety: Interview with highly anxious students. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 101–108). Prentice Hall.

Ran, C., Wang, Y., & Zhu, W. (2022). Comparison of foreign language anxiety based on four language skills in Chinese college students. *BMC Psychiatry*, *22*(558), 1–10. https://doi.org/10.1186/s12888-022-04201-w

Renko, K. (2012). *Finnish EFL learners' perceptions on errors, corrective feedback and foreign language anxiety* [Unpublished master's thesis]. University of Jyväskylä.

Russell, V. (2020). Language anxiety and the online learner. *Foreign Language Annals, 53*(2), 338–352. https://doi.org/10.1111/flan.12461

Saito, Y., & Samimy, K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate and advanced level college students of Japanese. *Foreign Language Annals, 29*, 239–251. https://doi.org/10.1111/j.1944-9720.1996.tb02330.x

Santos, A., Cenoz, J., & Gorter, D. (2017). Communicative anxiety in English as a third language, *International Journal of Bilingual Education and Bilingualism*, *20*(7), 823–836, https://doi.org/10.1080/13670050.2015.1105780

Sellers, V. (2000). Anxiety and reading comprehension in Spanish as a foreign language. *Foreign Language Annals, 33*, 512–521. https://doi.org/10.1111/j.1944-9720.2000.tb01995.x

Seltzer, K. A., & de los Rios, C. V. (2018). Translating theory to practice: Exploring teachers' raciolinguistic literacies in secondary English classrooms. *English Education, 51*(1), 49–79. http://works.bepress.com/kate-seltzer/5/

Shabani, M. B. (2012). Levels and sources of language anxiety and fear of negative evaluation among Iranian EFL learners. *Theory and Practice in Languages*, *2*(11), 2378–2383. https://doi.org/10.4304/tpls.2.11.2378-2383

Spradley, J. P. (1979). *The ethnographic interview*. Harcourt Brace Jovanich College Publisher.

Tan, K. H., Jospa, M. E., Mohd-Said, N. E., & Awang, M. M. (2021). Speak like a native English speaker or be judged: A scoping review. *International Journal of Environmental Research and Public Health, 18,* 1–16. https://doi.org/10.3390/ijerph182312754

Tanveer, M. (2007). *Investigation of the factors that cause language anxiety for ESL/EFL learners in learning speaking skills and the influence it casts on communication in the target language*. [Unpublished master's thesis]. University of Glasgow.

Tatar, S. (2005). Classroom participation by international students: The case of Turkish graduate students. *Journal of Studies in International Education, 9*(4), 337–355. https://doi.org/10.1177/102831530528096

Tercan, G., & Dikilitaş, K. (2015). EFL students' speaking anxiety: A case from tertiary level students. *ELT Research Journal, 4*(1), 16–27. https://dergipark.org.tr/en/pub/eltrj/issue/5485/74474

Thompson, A. S. & Khawaja, A. J. (2016) Foreign language anxiety in Turkey: The role of multilingualism. *Journal of Multilingual and Multicultural Development, 37*(2), 115-130, https://doi.org/10.1080/01434632.2015.1031136

Thompson, A. S., & Lee, J. (2013). Anxiety and EFL: Does multilingualism matter? *International Journal of Bilingual Education and Bilingualism, 16*(6), 730–749. https://doi.org/10.1080/13670050.2012.713322

Tobias, S. (1986). Anxiety and cognitive processing of instruction. In R. Schwarzer (Ed.), *Self-related cognition in anxiety and motivation* (pp. 35–54). Erlbaum.

Trang, T., Moni, K., & Baldauf, B. (2012). Foreign language anxiety and its effects on students' determination to study English: To abandon or not to abandon? *TESOL in context. Special Edition* S3, 1–14. https://tesol.org.au/wp-content/uploads/2019/10/TiC_S3_trang_tran.pdf

Tsang, A. (2020) The relationship between tertiary-level students' self-perceived presentation delivery and public speaking anxiety: A mixed methods study. *Assessment & Evaluation in Higher Education, 45*(7), 1060–1072. https://doi.org/10.1080/02602938.2020.1718601

Von Wörde, R. (2003). Students' perspectives on foreign language anxiety. *Journal of Linguistic Inquiry, 8*(1), 1–15. https://files.eric.ed.gov/fulltext/EJ876838.pdf

Wu, H., Garza, E., & Guzman, N. (2015). International student's challenge and adjustment to college. *Education Research International*, 1–9. https://www.hindawi.com/journals/edri/2015/202753/

Xianping, Z. (2003). *Language anxiety and its effects on oral performance in classroom.* [Unpublished paper]. Xiangfan University. https://www.researchgate.net/publication/264844780_Language_Anxiety_and_its_Effect_on_Oral_Performance_in_Classroom

Yan, X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language Learning, 58*(1), 151–183. https://doi.org/10.1111/j.1467–9922.2007.00437.x

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal, 75*(4), 426–439. https://doi.org/10.2307/329492

Young, D. J. (1992). Language anxiety from the foreign language specialists' perspective: Interview with Krashen, Omaggio Hadley, Terrell, and Rardin, *Foreign Language Annals, 25*, 157–172. https://eric.ed.gov/?id=ED335963

Zarei, N. & Moussavou, I. (2022). The impact of anxiety on ESL learners' English language fluency. *AGPE The Royal Gondwana Research Journal of History, Science, Economic, Political and Social Science*, *3*(5), 15–22. https://www.researchgate.net/publication/361360890

Zheng, Y. (2008). Anxiety and second/foreign language learning revisited. *Canadian Journal for New Scholars in Education/Revue canadienne des jeunes chercheures et chercheurs en education*, *1*(1), 1–12. https://files.eric.ed.gov/fulltext/ED506736.pdf

Ana Dili İngilizce Olan Öğrencilerin, Uluslararası Öğrencilerin Hissettikleri Olumsuz Değerlendirilme Korkusu ve Dil Kaygısındaki Rolü

Öz

Bu araştırma, lisansüstü düzeydeki uluslararası öğrencilerin, ana dili İngilizce olan öğrencilerin ağırlıkta olduğu sınıflardaki dil ve öğrenme deneyimlerini araştırmıştır. Çalışmanın amacı, ana dili İngilizce olan öğrencilerin varlığının, bir Amerikan üniversitesinde İngilizceyi ikinci dil olarak konuşan uluslararası öğrencilerin hissettikleri olumsuz değerlendirilme korkusu üzerindeki etkisini ortaya çıkarmaktır. Öncelikle 22 lisansüstü düzeydeki uluslararası öğrenciye bir anket uygulanmış ve ağırlıklı olarak ana dili İngilizce olan öğrencilerin bulunduğu sınıflarda olumsuz değerlendirilmeyle ilişkili yüksek kaygı düzeyleriyle öne çıkan beş uluslararası öğrenciyle röportaj yapılmıştır. Röportaj verileri içerik analizi ile analiz edilmiştir. Bulgular, uluslararası öğrencilerin olumsuz değerlendirilme korkusunun temel olarak eğitmenlerin uygulamalarına ve uluslararası öğrencilerin ana dili İngilizce olan akranlarının tutumlarına ilişkin algılarına bağlı olduğunu ortaya koymuştur. Dahası, artan olumsuz değerlendirilme korkusu, dil kullanımlarla ve derse genel katılımlarını olumsuz yönde etkilemiştir. Bulgular, lisansüstü programlarda daha kapsayıcı ve güvenli öğrenme ortamlarına ihtiyaç duyulduğunu göstermiştir.

Anahtar kelimeler: olumsuz değerlendirilme korkusu, ana dili İngilizce olan öğrenciler, dil kaygısı, dil kullanımı, derse katılım

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
2024, Vol. 41-3, No. 1
https://doi.org/10.52597/buje.1364919

The Analysis of Interjections in Two English Coursebooks

Zeynep Bütün Ikwuegbu[a]

Abstract

There has been a growing interest in the teaching of interjections in English, as interjections play an important role serving a bridge between verbal and non-verbal communication. However, interjections remain as an under-researched area in English language teaching (ELT). This study explores and compares the use and frequency of interjections in two English coursebooks. Yes You Can (written by non-native speakers of English and funded by the Turkish Ministry of Education) and Touchstone 2 (written by native speakers of English and published by Cambridge University Press) were chosen to collect data. A total of 97 dialogues were analysed. Although the findings indicated similarities between the two coursebooks regarding the functions and meanings of the interjections within the coursebooks, an obvious gap was observed between them regarding the diversity and amount of interjections integrated. The study offers suggestions to material developers, ELT teachers, language learners and authorities.

Keywords: interjections, ELT coursebooks, material evaluation, teaching spoken language

Article info

Received: 22.09.2023
Revised: 15.12.2023
Accepted: 13.01.2024
Published online: 22.04.2024

Introduction

This study investigates and compares two English coursebooks used in Türkiye, in order to understand the use and frequency of interjections in them. Interjections, defined as "words that constitute utterances by themselves and express a speaker's reaction towards an element in the socio linguistic context" (Ameka, 2006, p.4), were traditionally overlooked in English as they were considered as so-called "small words" (Petrova, 2020), or even `the outlaws of English` (Nordquist, 2012 as cited in Petrova, 2020). Therefore, whether interjections should be seen as a separate category in English has been very much controversial (Petrova, 2020), and they were attempted to be categorised under adverbs (Ashdowne, 2008), or even treated as paralinguistic phenomena (Wharton, 2003). Recently, however, there has been a growing interest to better understand the role of interjections in English. That may be linked to the fact that interjections are frequently used in everyday speech (Ponsonnet, 2023), and therefore play an essential role in communication and social interactions in English (Stivers, 2019).

In English language teaching (ELT), written language was seen as the main concern previously. However, this has shifted to spoken language, which is grounded on a more communicative approach (Brown & Yule, 1983). This has affected how ELT materials and curriculums should be designed, ideally integrating spoken language within them (Reber, 2011). Through the teaching of spoken language, learners can gain the opportunity to practise how to carry out daily conversations, both with native and non-native speakers of English (Elkilic & Genc, 2010).

Spoken language involves face-to-face interaction of two or more interlocutors in a shared space and time (Clark, 1998). When these interlocutors share common cultural and personal backgrounds, the resulting rich context allows speakers to reduce verbalisation to a minimum (Miller et al., 1998), through interjections (Goffman, 1981). However, interjections are culture specific, therefore there are often no similarities between interjections across languages (Wierzbicka, 1992), as a result of which, interjections can become obstacles to cross-cultural communication since interlocutors may not be familiar with them (Mao, 2017). This may lead to misunderstandings and/or communication breakdowns among interlocutors, which directly decreases the quality of communication (Hismanoglu, 2010). Therefore, learners should gain the ability to successfully use interjections

[a]*Kahramanmaraş Sütçü İmam Unıversity, Translation Studies, zeynep.butun89@gmail.com, ORCID: 0000-0001-5472-1178*

To cite: Bütün Ikwuegbu, Z. (2024). The analysis of interjections in two English coursebooks. *Boğaziçi Unıversity Journal of Education*, *41-3*(1), 19-34. https://doi.org/10.52597/buje.1364919

in English, as *the lingua franca*, to develop communicative competence (i.e., socio linguistic rules of a specific context), and authenticity in English (i.e., the proper use of spoken language in a specific context) (Celce-Murcia, 2008).

One effective way of achieving this is through English coursebooks (Sahragard et al., 2014), where interjections are integrated throughout to raise language learners` awareness to the natural flow of conversation (Reber, 2011), as well as to the contribution of interjections to communication (Cruz, 2009). This highlights the importance of the evaluation of teaching materials, which is carried out to identify the strengths and weaknesses of a course material, for a specific purpose (Slaouti et al., 2013). The proper choice of the coursebook for a programme, such as English coursebooks for teaching English, may have far reaching consequences, including spending on the module, learning experience, levels of proficiency, and ultimately, future employability of learners (Mishan & Timmins, 2015). However, as Hismanoglu (2010) mentions, there is a lack of developed materials teachers can utilise "to teach the communicative functions expressed by English interjections" (p.30). And, the teaching of interjections in English is still neglected (Ameka, 1992) and unfortunately under-researched (Petrova, 2020).

Previously, scholars mostly researched interjections in English, in relation to their linguistic features, such as pragmatic functions (e.g., Norrick, 2009). However, the existing literature lacks in providing a consistent categorisation of interjections (Stivers, 2019), and certainly more research is needed investigating the role of interjections in all stages of language teaching (Petrova, 2020). The few studies which analysed the coursebooks in terms of interjections (e.g., Reber, 2011; Sahragard et al., 2014) compared English coursebooks written by native speakers only – they will further be explained later under the section of relevant studies. This study differs from them by comparing an English coursebook written by Turkish authors, with one written by British authors, and focuses both on the forms and functions of interjections used in both coursebooks. The existing studies, although they looked at the forms of interjections, tended not to investigate the functions of interjections in detail. Also, in countries like Türkiye, where English is taught and learned as a foreign language, the teaching of interjections is often ignored by ELT teachers (Hismanoglu, 2010). And, therefore, learners` knowledgebase on interjections is generally very limited, as shown in Elkilic and Genc (2010), which can be improved by carefully designed and developed English coursebooks and teaching materials. This study attempts to fill these gaps in the literature by comparing two English coursebooks used to teach English in Türkiye, and evaluating them regarding the differences and similarities between them. Doing this will ultimately help to gain insights into the issues related to material development in Türkiye, and how to best help learners to increase their awareness of interjections, and enable them to implement interjections while communicating in English (Reber, 2011). The research questions are:

1. Which interjections are used and/or taught in *Yes You Can* and *Touchstone 2*?
2. What is the frequency of the interjections used and/or taught in *Yes You Can* and *Touchstone 2?*
3. How are *Yes You Can* and *Touchstone 2* different and similar in terms of the meanings and the functions of interjections used and/or taught?

Literature Review

This section briefly introduces two approaches to interjections, further focuses on the characteristics of interjections, presents a framework that informs this study, and also looks at the relevant studies in the literature.

From the sociolinguistic approach, interjections are not regarded as a part of the language (Goffman, 1981), as they deviate from observed morphological and syntactic patterns (Mushin et al., 2023). They are often considered as `response cries`, or a "ritualised act, in something like the ethological sense of that term" (Goffman, 1981, p.100). However, this is not to say interjections lack meaning and function, or does not have a role in social interactions (Mushin et al., 2023). Wilkins (1992), adopting the semantic approach, argue that "interjections are semantically rich and have a definite conceptual structure which can be explicated" (p.120). Therefore, they help speakers to reflect their state of mind, attitudes, and reactions towards situations (Ameka, 1992). Interjections convey messages without saying much (Wharton, 2003), which makes them the bridge between the verbal and non-verbal communication (Ameka, 2006). Somebody who says *Ugh!* or *Wow!*, for example, might show an immediate feeling of disgust or surprise, without actually saying *I am disgusted*, or *That`s amazing* (Goddard, 2014).

Characteristics of Interjections

Interjections cover a wide range of expressions, ranging from response cries such as *ouch* or *oops*, to *yeah* or *uh huh* (Stivers, 2019). Syntactically, interjections are independent and can stand alone as complete utterances, by, in most cases, signalling unexpected burst of emotions (Norrick, 2007). They can be used not only together with other patterns of the language, but also alone by helping to indicate speakers` emotional and mental state (Cruz, 2009). While in written texts they are separated from the rest of the sentence via a comma, a period or an exclamation mark (Wharton, 2009), a pause is given to separate them from other utterances in spoken language (Ameka, 1992). Although they are not productive since they do not take any derivation or inflection (Wilkins, 1992), interjections help the message to be conveyed during the conversation by acting as coded signals (Wharton, 2003). As such, they often have an illocutionary meaning as well, representing the implicit message intended by the speaker (Norrick, 2009). Some interjections have standardised meanings (Wilkins, 1992). However, many interjections convey attitudinal expressions depending on the linguistic and extra-linguistic contexts they are produced in (Denizot, 2013), or the intonation contour of the speaker (Norrick, 2009). As such, the meaning of interjections is co-constructed through the link between a sound pattern and a particular context (Reber, 2011).

As interjections appear in several formats, such as single words, sounds, phrases or utterances (Norrick, 2009), several attempts have been made to categorise them. For example, interjections have been classified with onomatopoeias; however, Meinard (2015) indicates that onomatopoeias are the transcription of any sound while interjections reflect the speakers` state of mind. And, interjections have largely been associated with expressing emotions (Reber, 2011), and specifically been linked to exclamative causes (Norrick, 2009). Yet, Ameka (1992), opposes to this by explaining that exclamations cover every utterance, and this categorisation, therefore, can lead to confusion. Others tended to categorise interjections with pragmatic markers (Norrick, 2007), and particles, including fillers, discourse markers and routines (Meinard, 2015). However, Ameka (1992), insists that interjections need to be considered as a distinct class that exists in all languages.

Ameka`s (1992) Typology of Interjections

Ameka (1992) talks about two main ways of classifying interjections: according to their forms and functions. In terms of their forms of, there are two kinds; namely, primary and secondary interjections (Ameka, 1992, 2006). Primary interjections are not used otherwise. They are defined as non-words since they do not fit in the rules of the language they belong to (Goffman, 1981), which is why they "tend to be phonologically and morphologically anomalous" (Ameka, 1992, p.105). *Oh, mhm, huh, wow, ohh, oops, ouch, mm, eh, ah, psst* are some examples of primary interjections (Wharton, 2009). On the other hand, secondary interjections are "forms that belong to other word classes based on their semantics and are interjections only because they can occur by themselves non-elliptically as one-word utterances" (Ameka, 1992, p.105). Secondary interjections tend to be lexical items carrying a concrete semantic meaning (Ameka, 1992), and they are morphologically and semantically motivated (Denizot, 2013). Unlike primary interjections, which act as turn initiators, secondary interjections display a range of functions (Norrick, 2009). Additionally, secondary interjections have an independent semantic value, and they can be used to express a mental attitude, or state. Most of the secondary interjections consist of one morpheme (which is called monomorphemic) such as *yeah!, God!, Gosh!, boy!*; however, there are also some examples of multi-morphemic interjections such as *Goddamit!* (Ameka, 2006).

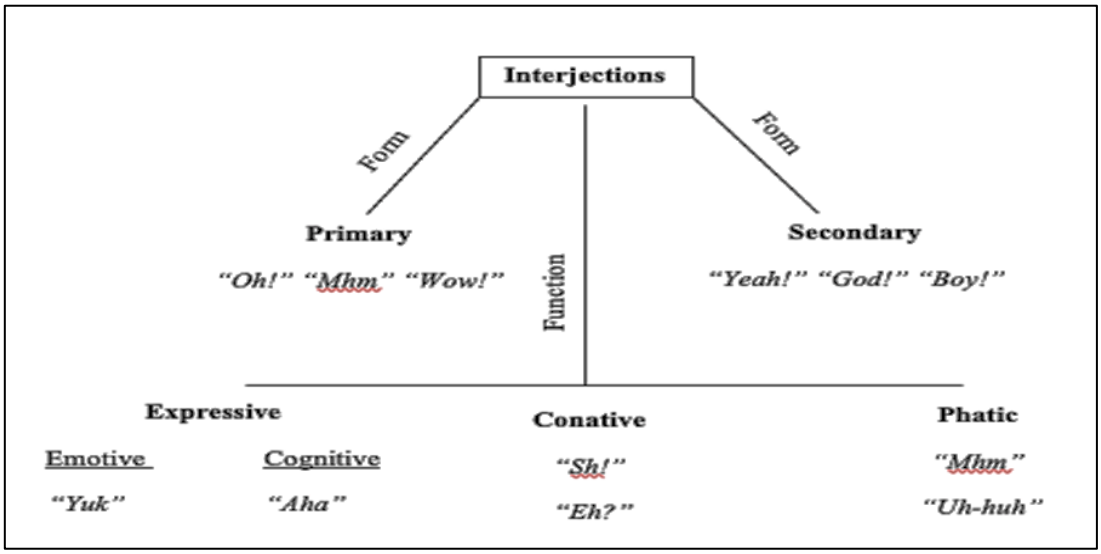
Regarding their function, Ameka (1992) classifies interjections as three; namely, expressive, conative, and phatic interjections (Ameka, 1992). Expressive interjections are the vocal gestures that reflect the mental state of the speaker and are subdivided into two (Ameka, 1992). Emotive interjections, as the first group of expressive interjections, reflect speakers` emotions and sensations that they have at the time of speech (Ameka, 1992), and they have the component of *I feel something* in their meaning (Wierzbicka, 1992). *Yuk* for disgust, *wow* for surprise, and *ow* for pain are some examples that can be given for this type of interjections (Wilkins, 1992). Cognitive interjections, as the second group of expressive interjections, reflect speakers` thoughts and state of knowledge at the time of speech (Ameka, 1992). This type has the component of *I think something*, or *I know something* in their meaning (Wierzbicka, 1992). For example, *aha* might be used to mean *I understand* (Wierzbicka, 1992), and *ehe* might mean *I now remember* (Ameka, 1992). Conative interjections can be claimed to be listener-oriented since they are the expressions which are directed to listeners to get their attention, response, or reaction (Ameka, 1992), and they often mean *I want something* (Wierzbicka, 1992). For example, *sh!* might be used for *I want silence here*, and *eh?* for *I want to learn something* (Ameka, 1992). As the third group, phatic interjections show the mental attitude of speakers towards an on-going conversation (Ameka, 1992). They can be

used to establish a conversation, and since they serve as feedback, they can be used to maintain the conversation as well (Celce-Murcia & Olshtain, 2000). *Mhm*, *uh-huh*, and *yeah* can be given as examples for phatic interjections (Ameka, 1992).

In this study, I utilised Ameka`s (1992) typology, as it is succinct and sufficiently explanatory (Nakatani, 2005). Doing this helped to answer the research questions, as it investigates both forms, and functions of the interjections. Based on Ameka (1992), Figure 1 was depicted, to guide the analysis of the interjections.

Figure 1

Interjections in Relation to Their Forms and Functions (based on Ameka, 1992)



Relevant Studies

As indicated previously, interjections is a poorly researched area (Petrova, 2020). Some studies researched interjections in English as a foreign language (EFL) contexts, without a specific focus on the teaching of interjections. Thompson (2022), for example, investigated interjections in one EFL corpus in Asia, in comparison to a non-EFL corpus, and found fewer interjections with less variety in the EFL corpus, which, according to the researcher, might have been linked to EFL users` lack of knowledge on certain interjections. Other studies mostly looked at language learners` capabilities of using interjections. For example, Hismanoglu (2010) investigated EFL learners` competence in using the appropriate interjections in Cyprus, and found that although learners could use the interjections appropriately at times when those interjections had a similar meaning and function to the ones in the learners` native language (L1), they failed in using other interjections which differed from the ones in their L1. The researcher, therefore, underlines the importance of exposing language learners to authentic real-life materials and situations. Similarly, another study was conducted in Türkiye with 42 intermediate-, and 40 advanced-level EFL learners, and found that learners, regardless of their level, were familiar only with the very commonly used interjections (Elkilic & Genc, 2010). The researchers conclude that English coursebooks need to be prepared by professionals, and learners should be explicitly taught about the differences between Turkish and English interjections (Elkilic & Genc, 2010). In Japan, Talandis and Stout (2015), despite not focusing particularly on interjections, conducted action research and designed a curriculum by means of personalised topics, direct teaching of pragmatics, and frequent assessment of oral skills to improve EFL learners` speaking skills. They found that explicit teaching of conversation strategies helped learners to interpret the meanings of interjections and solve interactional difficulties (Talandis & Scout, 2015). What these studies seem to have in common is the implication that explicit teaching of interjections, through real-life materials, can help language learners to better learn and use interjections while communicating. This brings about the importance of English coursebooks in the teaching of interjections; however, there is very little research focusing on the integration of interjections in English coursebooks. In one study, Reber (2011) carried out a survey of English coursebooks used in German elementary and secondary schools, and found that interjections were included in constructed conversation examples in the coursebooks. However, there was a lack of instruction regarding the meaning and use of interjections, and, a lack of information regarding visual actions in the coursebooks, which, according to Reber

(2011), might limit learners from inferring the correct meaning of interjections. In another study, Sahragard et al. (2014) investigated three coursebooks in terms of the frequency of interjections used in them, and one coursebook was found integrating more interjections with a greater variety. The researchers highlight the necessity of developing more authentic (from real-life situations) materials.

Methodology

A qualitative approach was adopted for this study as the purpose was to analyse two English coursebooks; namely, *Touchstone 2* (McCarthy et al., 2012) and *Yes You Can* (A.2.2 and A.2.3, as a set of two used in sequence) (Baydar Ertopcu et al., 2015; Persembe et al., 2015), to gain an understanding of the forms and functions of interjections used in them (Gray, 2014). Purposive sampling, specifically, criterion sampling was employed, as the focus of the study was on interjections within the chosen coursebooks (Gray, 2014). Therefore, all pages in both coursebooks were reviewed, a total of 402 pages constituted the unit of analysis, with all written dialogues and listening transcripts in both coursebooks, which were the sections including interjections (Patton, 2015). Table 1 shows the number of units, pages and dialogues in both coursebooks.

Table 1

Number of Units, Pages and Dialogues in the Coursebooks

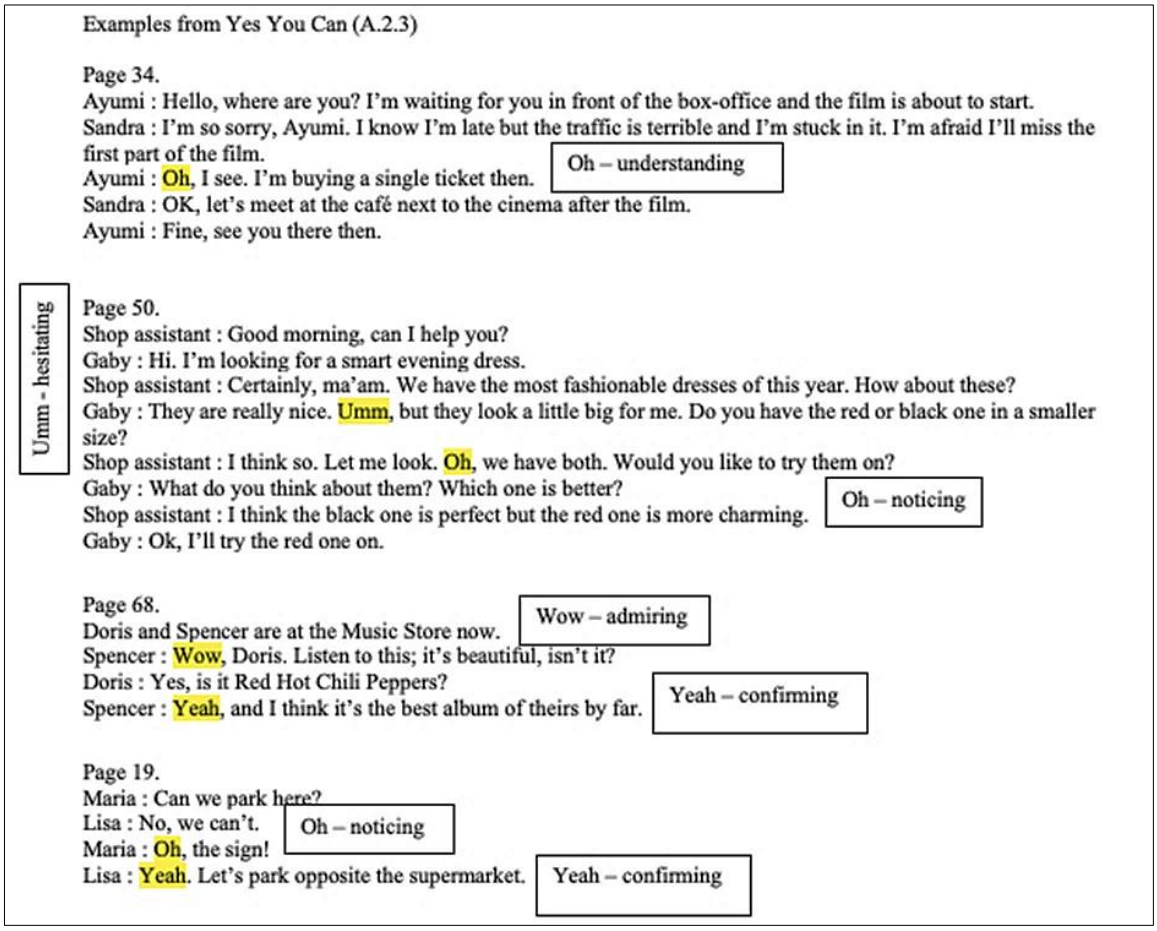
		Yes You Can	Touchstone 2	TOTAL
Units		16	12	28
Pages		244	158	402
Dialogues	Written	25	35	60
	Spoken	17	20	37

The coursebooks represented around the same level of proficiency, namely, intermediate level. *Yes You Can* was developed by Turkish writers and funded by the Turkish Ministry of Education, while *Touchstone 2* was written by British writers and published by Cambridge Press Publications. *Touchstone 2* was chosen purposefully for this comparison since it was designed according to the spoken corpora as stated in the coursebook. According to Mishan (2005), spoken corpora is an incomparable resource for coursebook writers as it offers `pure` language loaded with contextual features. Comparing *Touchstone 2* with *Yes You Can* (in which corpora was not taken into consideration) provided insights regarding the process of material development.

Content analysis was carried out on the gathered data, which helped to make evaluative comparisons between two coursebooks, according to the established goal, i.e., identifying the forms and functions of interjections (Drisko & Maschi, 2016). Throughout the analysis, these steps were followed; selecting the text to be examined, specifying the unit of analysis, determining the category scheme, selecting final categories and analysing and interpreting findings (Gray, 2014). The data was coded both deductively and inductively (Drisko & Maschi, 2016). The deductive part was informed by the literature, about the forms of interjections (Drisko & Maschi, 2016). To answer the first research question, Ameka`s (1992) classification of interjections was employed, to detect the interjections used and/or taught in *Yes You Can* and *Touchstone 2*. In order to answer the second research question, the number and the frequency of interjections was calculated manually and determined for both coursebooks. As the aim was to understand the meaning and the functions of interjections, discourse analysis, in addition to content analysis, was carried out for the third research question, which was when the data was analysed more inductively by reading and rereading the unit of analysis for multiple times. The focus of the discourse analysis is "on the use of particular words, phrases, idioms, smiles, metaphors, kinds of rhetoric, and so on" (Thomas, 2013, p.242), which in this case was interjections used and/or taught in the coursebooks. This part of the analysis included interpreting interjections by taking into consideration the dialogue so that which function and meaning each interjection was signalling to could be found out (Gray, 2014). Figure 2 shows an example of how the functions of interjections were determined.

Figure 2

Examples from Yes You Can, in Terms of the Coding Process



This interpretation stage was followed by grouping and categorising the functions and the meanings of each interjection. In addition, Oxford English Dictionary Online (2012) was consulted to check whether or not other alternative meanings and usages of each observed interjection existed. As the last step, *Yes You Can* and *Touchstone 2* were compared to each other to find the similarities and differences. I was involved in the data analysis process; however, to ensure the quality of the coding stage, I sought intercoder reliability on a small set of data. This ensured that the data was categorised in the same way, and the created codes were consistent (Drisko & Machi, 2016). Additionally, I constantly assessed the codes over time and made amendments if necessary, to make sure of the internal consistency of the coding process, which contributed to the intracoder reliability of the coding process (Neuendorf, 2017). Doing these and providing examples from the data, to support the inferences and conclusions drawn, contributed to the validity of the coding process as well (Neuendorf, 2017).

Findings

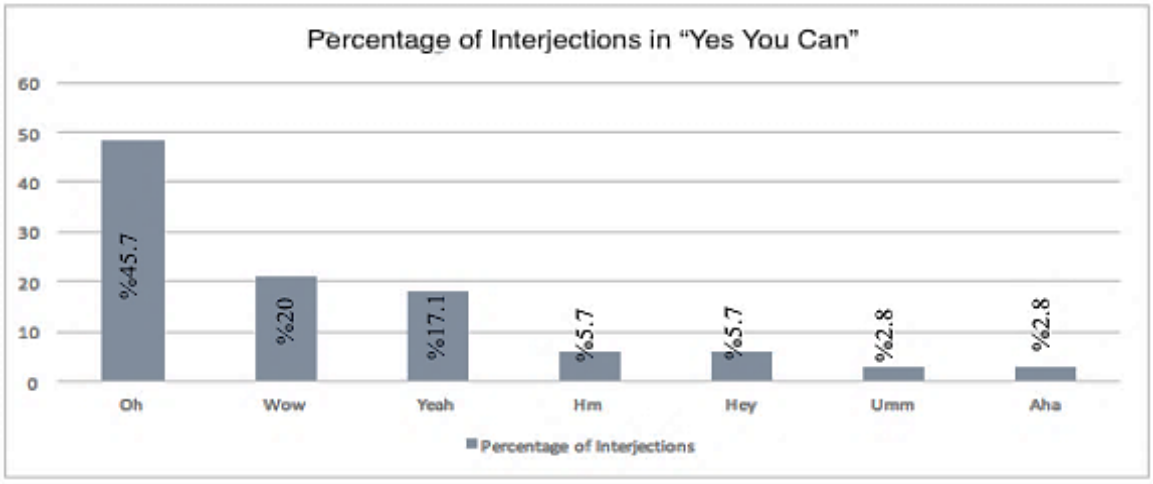
The findings of this study will be presented in accordance with the research questions. While the first and the second research questions will be answered together as they are linked to each other, the third research question will be explained separately.

Comparison of Yes You Can and Touchstone 2 Regarding the Interjections Used within Them, and Their Frequency

Each dialogue in both coursebooks was studied carefully in order to detect the interjections. Based on the analysis of 42 dialogues, *Yes You Can* was found to have 35 interjections with 7 different types. *Oh* and *wow* were the most frequent interjections in the coursebook while *umm* and *aha* were found to be the least frequent ones. Figure 3 indicates the percentages of the interjections in *Yes You Can*.

Figure 3

Percentage of Interjections in Yes You Can

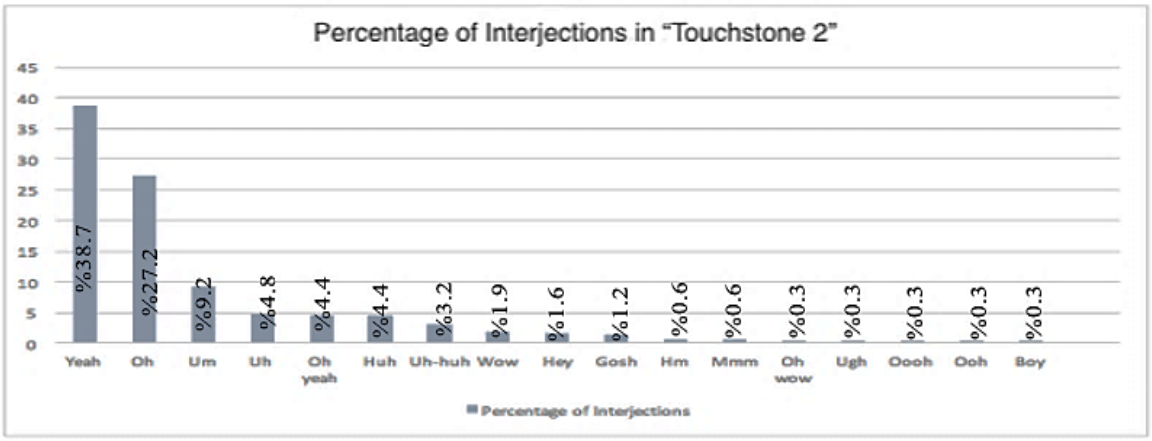


As Figure 2 shows, all interjections were structured either as one word or as non-word. While six types of interjections, which corresponds to 85.71% of all interjections found in *Yes You Can*, were primary with the range of *oh*, *wow*, *hm*, *hey*, *umm*, and *aha*; *yeah* was the only example for secondary interjection, which corresponds to 14.28%. Moreover, after the analysis of the coursebook, it was determined that no explicit instruction regarding the meaning or the usage of interjections was provided.

After analysing 55 dialogues, *Touchstone 2* was found to include 312 interjections with 17 different types. While *yeah* and *oh* were found to be the most frequently repeating interjections, *oh wow*, *ugh*, *ooh* and *boy* were found the be the least frequently repeating interjections. Figure 4 demonstrates the percentages of the interjections in *Touchstone 2*.

Figure 4

Percentage of Interjections in Touchstone 2



As can be inferred from Figure 3, 15 out of 17 types of interjections, which corresponds to 88.23% of all interjections found in *Touchstone 2*, were structured as either one word or non-word such as *yeah*, *oh*, *um*, *uh*, *huh*, *uh-huh*, *wow*, *hey*, *Gosh*, *hm*, *mmm*, *ugh*, *oooh*, *ooh* and *boy*. *Oh yeah* and *oh wow*, which corresponds to 11.76%, were found to be interjections that were structured as two-words. Moreover, based on the analysis, it can be indicated that 13 types of interjections (*oh, um, uh, huh, uh-huh, wow, hey, hm, mmm, oh wow, ugh, oooh, ooh*), which corresponds to 76.4%, were primary while four of them (*yeah, oh yeah, Gosh, boy*), which corresponds to 23.5%, were secondary. Unlike *Yes You Can*, instructions regarding the function and meaning of the interjections were found in *Touchstone 2*. For example, in Unit 3 in *Touchstone 2*, a glossary section is provided, and *oh*, *wow*, *oh wow*, *Gosh* are introduced as expressions to show surprise in informal conversations (p.27). Also, on the same

page, the frequency of these interjections in the spoken corpora is provided, to show learners which ones are more frequently used.

By comparing Figure 3 and Figure 4, differences and similarities between the coursebooks regarding the diversity of the interjections could be understood. *Yeah*, *oh*, *wow* and *hey* were the interjections that were used in both coursebooks. However, while *aha* was used only in *Yes You Can*, *uh*, *oh yeah, huh, uh-huh, Gosh, mmm, oh wow, ugh* and *boy* were used only in *Touchstone 2*. Moreover, some variants of the interjections were found in the coursebooks. While *hm, umm*, and *oh* were used in *Yes You Can*, *hmm, um, oooh* and *ooh* were used in *Touchstone 2*.

Comparison of Yes You Can and Touchstone 2 Regarding the Meanings and Functions of Interjections Used within Them

In addition to the diversity of the interjections used in the coursebooks, similarities and differences regarding their functions and meanings were detected between the coursebooks (see Appendix A and Appendix B for a detailed list of interjections with regards to their functions, meanings, frequency, and percentages). *Oh* was one of the largely used interjections in both *Yes You Can* and *Touchstone 2* as shown in Figure 3 and Figure 4. Table 2 shows the functions and meanings of *oh* in both coursebooks.

Table 2

Functions and Meanings of "Oh" in Both Coursebooks

Functions	Meanings
Yes You Can	
To show a change in cognitive process	"I see/notice/remember"
To show turn-taking	No precise meaning
To show affection	"It is bad!"
To show disappointment	"I am disappointed"
To show pleasure	"Great!"
To show disagreement	"No!"
To show annoyance	"I am angry"
Touchstone 2	
To show turn-taking	No precise meaning
To show a change in cognitive process	"I see/notice/remember"
To show surprise	"Interesting"
To show disappointment	"I am disappointed"
To show affection	"I like it"
To show affection	"It is bad"
To show annoyance	"I am angry"

As shown, *oh* was used for several functions and meanings in both coursebooks. Being used as a cognitive interjection by speakers, *oh* indicated a change in cognitive process with a percentage of 43.75 in *Yes You Can* and with a percentage of 32.94 in *Touchstone 2*. In this case, speakers used *oh* to show an understanding, noticing and remembering moment throughout the conversation. An example from *Touchstone 2* is:

Example 1

Alicia: It`s Mom`s birthday on the first. Remember? She is going to be 50!

Dave: Oh, that`s right. What are you going to get her?

Oh was used as a phatic interjection as well where the listeners reacted to a remark or to a question to imply that they are taking the turn. The percentage for this function was 18.75 in *Yes You Can*, and 43.52 in *Touchstone 2*. In addition, it was used as an emotive interjection to show a range of emotions. In both coursebooks, it was used to show affection; however, while in *Yes You Can* the meaning implied by the speaker was *it is bad*, in *Touchstone 2* the speaker showed affection to convey two meanings; *it is bad* and *I like it*. Moreover, in both coursebooks *oh* was used to reflect both disappointment and anger. An example from *Yes You Can* is:

Example 2

Doris: How are you going to pay for it?

Spencer: Good point! I don`t have my credit card with me. Can I use yours?

Doris: Oh, Spencer! You are hopeless!

In addition to the abovementioned similarities between *Yes You Can* and *Touchstone 2*, regarding the functions and meanings derived from the use of *oh*, the coursebooks had differences as well. For example, *oh* was used to show surprise in *Touchstone 2* for nine times (10.58%), and it was used to signal pleasure and disagreement with the meanings of *great* and *no* in *Yes You Can*.

Another similarity between the coursebooks was the way *wow* was used (see Appendices A and B). *Wow* was used as an emotive interjection in order to show admiration and surprise. While it was used for five times (71.42%) with the meaning of *amazing* in *Yes You Can*, it was used for two times (33.33%) with the same meaning in *Touchstone 2*. Additionally, in both coursebooks, speakers used *wow* to show their surprise. However, the percentage was different as it was used for four times (66.66%) in *Touchstone 2* when the frequency was two (28.57%) in *Yes You Can*. An example from *Touchstone 2* is as follows;

Example 3

Man: I have the same dream every night.

Woman: Every night? Wow! What do you dream about?

Another widely used interjection in both coursebooks was *yeah*. Table 3 indicates the functions and meanings of yeah in both coursebooks.

Table 3

Functions and Meanings of "Yeah" in Both Coursebooks

Functions	Meanings
Yes You Can	
To show confirmation	"Yes"
To show agreement	"You are right"
Touchstone 2	
To show confirmation	"Yes"
To show agreement	"You are right"
To give feedback	"I am listening"
To show turn-taking	No precise meaning
To ask for confirmation	"Is it so?"
To show hesitation	"I am not sure"
To show self-confirmation	"I am sure"

In order to maintain the conversation, *yeah* was used in both coursebooks as a phatic interjection. While it was used for five times (83.33%) in *Yes You Can* with the meaning of yes for showing confirmation, the frequency was 59 (48.76%) in *Touchstone 2*. An example for this meaning from *Yes You Can* is as follows;

Example 4

Todd: Hey, Stewart, have you heard about the Van Gogh Alive exhibition?

Stewart: Yeah, I`ve already heard the name but I don`t have any idea about the details. Where is it?

In addition to this, *yeah* was used by speakers to show that they agreed with their interlocutors, in which case it meant *you are right*. With this meaning, it was used for once (16.16%) in *Yes You Can*, while the frequency was 38 (31.40%) in *Touchstone 2*.

Example 5, Touchstone 2

Chris: You know, we should take a few days off sometime.

Adam: Yeah, we should. Definitely.

Apart from these functions, *yeah* was used in *Touchstone 2* for several other functions as well such as back channelling the speaker, asking for confirmation, self-correction, showing hesitation and turn-taking. In *Yes You Can*, however, the function of *yeah* was limited to showing confirmation and agreement.

In both coursebooks, *hey* was used as a conative interjection to draw the listener`s attention (see Appendices A and B). While it was used for two times (100%) in *Yes You Can*, in *Touchstone 2* it was used for four times (80%) for this function.

Example 6, Yes You Can

Todd: Hey, Stewart, have you heard about the Van Gogh Alive exhibition?

Stewart: Yeah…

Hey was also used as an emotive interjection in *Touchstone 2* by the speaker to show her anger; however, in *Yes You Can*, this function was not given place. With the purpose of showing speakers` hesitation, *hmm* and *hm* - as a variant of *hmm* - were used for two times (100%) in *Touchstone 2* and two times (100%) in *Yes You Can* respectively (see Appendices A and B). An example from *Touchstone 2* is:

Example 7

Ray: And there`s good food. You can get all kinds of tacos and things. Do you want to go?

Tina: Hmm. Well, maybe.

As the last common interjection between the coursebooks, *um* and *umm* were used in the coursebooks to show speakers` hesitation (see Appendices A and B). In this case, the meaning attributed to *um* and *umm* was *I am not sure*. In *Touchstone 2*, *um* was used for seven times (24.13%) with this meaning while *umm* was used with the same meaning for one time (100%) in *Yes You Can*.

Example 8, Touchstone 2

Mark: I got a new apartment. It is really nice, but like I said, I can`t find my TV, I mean, it`s in one of those boxes, so, uh…so, could I borrow yours? Just for a couple of days?

Jenny: Um, I guess. Like when?

In addition to this function of *um*, it was also used to maintain the conversation as a phatic interjection by indicating speakers` thinking process and their intention to get to another topic in *Touchstone 2*. Even though the coursebooks seemed to have similarities in terms of the functions and meanings of the interjections mentioned thus far, they differed from each other with regards to the choice of interjections they used to indicate a change in cognitive process. In addition to *oh*, which was used in both coursebooks for this purpose and explained above, aha seemed to be used for this purpose in *Yes You Can*, while in *Touchstone 2* it was *oh yeah* and *ooh*.

Discussion

The findings showed that *Yes You Can* and *Touchstone 2* were similar in terms of integrating Ameka`s (1992) primary and secondary interjections; and they both used primary interjections more commonly than secondary interjections. However, an obvious gap was observed between the coursebooks. *Touchstone 2* was found to include a wider range of interjections with more frequent use compared to *Yes You Can* in which both the variety and frequency of the interjections was quite limited. In addition, while *Touchstone 2* seemed to introduce multiple meanings and functions for most of the interjections used in the coursebook, Yes You Can seemed to fail in achieving this. And, *Touchstone 2* introduced a variety of interjections for the same meaning/function, such as *yeah, oh yeah, huh* for confirming; *yeah, oh yeah* for agreeing; *yeah, uh-huh* for giving feedback; *yeah, oh, uh-*

huh for turn-taking; *yeah, um, uh, hmm* for hesitating; *oh, oh yeah, ooh* for showing a cognitive change; *oh, huh, wow, Gosh, boy* for showing surprise; *oh, hey* for showing annoyance; *um, uh* for showing a thinking process; *um, uh* for transitioning to the (main) point; and *wow* and *oh wow* for admiring (see Appendices A and B). *Yes You Can*, on the other hand, provided very few examples of the interjections which were used for the same meaning/function; and these were *umm* and *hm* for hesitating, and *oh* and *aha* for showing a cognitive process. Learners need to know that there are several types of interjections, serving different purposes such as expressing emotions or processing a thought (Cruz, 2010). Therefore, their being introduced to as many different interjections as possible appears to be very important, even if not every interjection is used as often in everyday conversations (Aijmer, 2004). That is because, learning how to use interjections while communicating can give learners a sense of authenticity and expressiveness, helping them to develop their communicative competence (Petrova, 2020). The finding regarding *Touchstone 2* seems to align with Sahragard et al. (2014), in which the researchers compared three different coursebooks (*Interchange 3, Top Notch 3 A-B, Touchstone 4*), all written by native speakers, and found that *Touchstone 4* integrated a greater diversity and higher frequency of interjections compared to the other two. This difference, although not specified by the researchers, might result from whether the coursebooks were designed by taking the spoken corpora into consideration. Accordingly, the researchers suggest that material developers should design authentic coursebooks by integrating interjections more effectively (Sahragard et al., 2014).

Given the number and variety of the interjections used in the coursebooks, *Yes You Can* obviously lacks in providing learners with a wide range of interjections since the amount of the interjections found in *Touchstone 2* is almost nine times more than the number of the interjections *Yes You Can*. A possible reason for this gap may be linked to the fact that *Touchstone 2* was designed by using the spoken corpora as explicitly stated in the coursebook. As such, it may not be surprising that *Touchstone 2* included more genuine communications between people from real life, rather than prepared communications in artificial conditions (Mishan, 2005). In the development of *Yes You Can*, however, corpora were not taken into consideration. This looks in alignment with Paksoy and Harmaoglu (2017). The researchers investigated the authenticity of the language in four English coursebooks used in Turkish high schools, and did not specifically look at interjections. However, after comparing the coursebooks with the British National Corpus, Paksoy and Harmaoglu (2017) observed that the coursebooks had little similarity to the authentic language, which could limit the opportunities for language learners to be exposed to real-life language.

Yes You Can and *Touchstone 2* differed in terms of the number of the dialogues as well, which implies that *Touchstone 2* gives more emphasis to the spoken language than *Yes You Can* does. The analysis of the structure of the dialogues showed that the dialogues in *Touchstone 2* reflect the real-life conversation better than the dialogues integrated in *Yes You Can*. And, many of the dialogues in *Yes You Can* are structured as interviews. Accordingly, because of the nature of the dialogues (formal vs informal), the more informal real-life dialogues (emphasising the spoken language) were found to include more interjections than the formal interview type dialogues (emphasising the written language). According to Cutting (2008), interviews tend to include unequal balance among the interlocutors since the interviewer always asks and the interviewee always explain. However, the main purpose of the dialogues should be to promote interaction by providing several characteristics of the spoken language rather than merely talking (Cutting, 2008).

Learners were provided with instructions in *Touchstone 2* in terms of the use of *oh, wow, oh wow, Gosh, oh my Gosh* to show surprise, for example. However, no explicit instruction was provided in *Yes You Can* regarding the meaning or the function of interjections. This finding seems to align with Reber (2011), which showed, after analysing three coursebooks used to teach English in Germany, that the coursebooks did not offer accurate and detailed information for the meanings of the interjections that were taught. Explicit teaching of interjections is considered essential as interjections help learners to express their attitudes, emotions and values, all of which influence the flow of the conversation (Reber, 2011), and determine whether learners become a competent speaker of the target language (Cruz, 2010). In the absence of instructions regarding the use and meaning of interjections, as Reber (2011) indicates, learners interpret and incorporate them in their learning process, which may lead to misconstrual.

Conclusion

This study investigated and compared *Yes You Can* and *Touchstone 2* regarding the meanings and frequency of interjections, with the purpose of answering three research questions. While the coursebooks were found to be similar in terms of integrating more of primary interjections, rather than secondary interjections, and regarding

the functions and meanings attributed to these interjections within both coursebooks, there were differences as well. *Touchstone 2* integrated a wider range of interjections, with higher frequency, and introduced several interjections for the same meaning/function. Also, while *Yes You Can* did not provide any explicit instruction regarding the function or meaning of interjections, *Touchstone 2* did.

This study adopted a qualitative interpretivist approach and therefore the findings cannot be generalised, and only apply to the analysed coursebooks. Still, it provides insights into how to teach interjections and also how to develop better coursebooks for teaching interjections in English. The findings suggest that integrating everyday dialogues into the coursebooks, as in *Touchstone 2*, meant to include more interjections since they reflected more of informal and spoken language. Additionally, benefitting from the spoken corpora, at the stage of developing an English coursebook, can help to understand the variety of interjections used in daily conversations and the meanings/functions of them. Moreover, explicit teaching of interjections can help to draw language learners` attention to the importance of using interjections while communicating in English.

Several stakeholders can benefit from the findings of this study. Firstly, coursebook writers need to take the spoken corpora into consideration at the stage of developing English coursebooks, to provide more authentic and real-life communication situations, and also to focus more on frequently used language. Doing this especially important in contexts where English is not spoken as the first language, as in Türkiye, as coursebook writers` intuitive knowledge regarding the use of English can be misleading, or even wrong (Paksoy & Harmaoglu, 2017). Coursebook writers ought to provide information regarding the functions and meanings of interjections in a glossary section, where the meaning/function of an interjection is difficult to be inferred from the dialogue itself (Reber, 2011). ELT teachers need to make better decisions on the coursebooks they use for their students. Also, when the coursebooks do not offer authentic enough language, they should utilise additional resources to explicitly teach interjections. Learners need to realise the role of interjections to become more competent in English, and try to use interjections in their conversations. Lastly, authorities ought to provide trainings to ELT teachers, to raise their awareness on interjections. This is especially important for non-native ELT teachers, as they may lack the mastery of the usage of interjections (Cruz, 2010). And, authorities should prioritise the teaching of spoken language, and accordingly choose the coursebook writers.

Code of Ethics

All information in this paper has been obtained and presented in accordance with academic rules and ethical concerns. The study did not require an ethical approval as it does not involve human or animal participants.

Statement of Interest

I declare that there is no conflict of interest.

References

Aijmer, K. (2004). Interjections in a contrastive perspective. In E. Weigand (Ed.), *Emotion in dialogic interaction: Advances in the complex* (pp. 99–120). John Benjamins Publishing Company.

Ameka, F. (1992). Interjections: The universal yet neglected part of speech. *Journal of Pragmatics, 18*, 101–118. https://doi.org/10.1016/0378-2166(92)90048-G

Ameka, F. (2006). Interjections. In K. Brown (Ed.), *Encyclopaedia of language and linguistics* (pp.743–746). Elsevier.

Ashdowne, R. (2008). Interjections and the prts of speech in the ancient grammarians. *Henry Sweet Society for the History of Linguistic Ideas Bulletin*, *50*(1), 3–15.

Baydar Ertopcu, F., Inci, H., & Ozbicakci Samur, S. (2015). *Yes You Can: Student`s Book A2.2* (4th ed.). Devlet Kitaplari.

Brown, G., & Yule, G. (1983). *Teaching the spoken language: An approach based on the analysis of conversational English*. Cambridge University Press.

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. S. Jorda (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.

Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching*. Cambridge University Press.

Clark, A. (1998). Magic words: how language augments human computation. In P. Carruthers & J. Boucher (Eds.) *Language and thought: interdisciplinary themes* (pp. 21–39). Routledge.

Cutting, J. (2008). *Pragmatics and discourse: A resource book for students* (2nd ed.). Routledge.

Cruz, M. P. (2009). Towards an alternative relevance –theoretic approach to interjections. *International Review of Pragmatics, 1*, 182–206. https://doi.org/10.1163/187731009X455884

Cruz, M. P. (2010). Teaching interjections in the ESL/EFL class: A pragmatic approach. *Estudios de Metodologla de la Lengua Inglesa, 5*, 23–33.

Denizot, C. (2013). Interjections. In G. K. Giannakis (Ed.), *Encyclopaedia of ancient Greek language and linguistics*. Brill.

Drisko, J. W., & Maschi, T. (2016). *Content analysis*: *Pocket guides to social work research methods*. Oxford University Press.

Elkilic, G., & Genc, B. (2010). *Turkish EFL Students Failure of the Use of Interjections Properly: The Example of Kafkas University*. 2nd International Symposium on Sustainable Development.

Goddard, C. (2014). Interjections and emotion (with special reference to "surprise" and "disgust"). *Emotion Review, 6*(1), 53–63.

Goffman, E. (1981). *Forms of talk.* Basil Blackwell.

Gray, D. E. (2014). *Doing research in the real world* (3rd ed.). SAGE Publications Ltd.

Hismanoglu, M. (2010). Interjections in English: Neglected but important aspect of foreign language learning and teaching. *Journal of Theory and Practice in Education, 6*(1), 17–35.

Mao, A. M. (2017). Conceptuality and context-sensitivity of emotive interjections. *Open Journal of Modern Linguistics, 7*(1), 41–51. DOI: 10.4236/ojml.2017.71004

Meinard, M. E. M. (2015). Distinguishing onomatopoeias from interjections. *Journal of Pragmatics, 76*, 150–168. https://doi.org/10.1016/j.pragma.2014.11.011

Miller, J. E., Miller, J., & Weinert, R. (1998). *Spontaneous spoken language: Syntax and discourse*. Oxford University Press.

Mishan, F. (2005). *Designing authenticity into language learning materials*. Intellect Books.

Mishan, F., & Timmis, I. (2015). *Materials development for TESOL*. Edinburgh University Press.

Mushin, I., Blythe, J., Dahmen, J., De Dear, C., Gardner, R., Possemato, F., & Stirling, L. (2023). Towards an interactional grammar of interjections: Expressing compassion in four Australian languages. *Australian Journal of Linguistics, 43*(2), 158–189.

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication ctrategy use. *The Modern Language Journal, 89*(1), 76–91. http://www.jstor.org/stable/3588552

Neuendorf, K. A. (2017). *The content analysis guidebook*. SAGE Publications.

Norrick, N. R. (2007). Discussion article: Pragmatic markers, interjections and discourse. *Catalan Journal of Linguistics, 6*, 159–168. https://raco.cat/index.php/CatalanJournal/article/view/74215

Norrick, N. R. (2009). Interjections as pragmatic markers. *Journal of Pragmatics. 41*, 866–891. https://doi.org/10.1016/j.pragma.2008.08.005

Oxford Dictionaries (2012). The world`s most trusted dictionaries. Oxford University Press.

Paksoy, E., & Harmaoglu, O. (2017). Corpus based authenticity analysis of language teaching course Books. *International Journal of Languages` Education and Teaching, 5*(4), 287–307. http://dx.doi.org/10.18298/ijlet.2324

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). SAGE Publications.

Persembe, E., Bulug, N., & Eroglu Canmetin, Z. Z. (2015). *Yes you can: Student`s book A2*.3 (4th ed.). Devlet Kitaplari.

Petrova, K. (2020). Interjections and L2 learning and teaching. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, *1*, 323–336.

Ponsonnet, M. (2023). Interjections. In C. Bowern (Ed.), *The Oxford guide to Australian languages*, (pp. 564–572). Oxford University Press.

Reber, E. (2011). Interjections in the EFL classroom: teaching sounds and sequences. *ELT Journal*, *65*(4), 365–375. https://doi.org/10.1093/elt/ccq070

Sahragard, R., Mavaddat, R., Bayani, S. R., & Safavian, M. (2014). Investigating the implementation of interjections in three current EFL coursebooks. *Procedia- Social and Behavioural Sciences*, *98*, 1621–1630. https://doi.org/10.1016/j.sbspro.2014.03.586

Slaouti, D., McDonough, J., Shaw, C., & Masuhara, H. (2013). Technology in ELT. In J. McDonough, C. Shaw, H. Masuraha (Eds.), *Materials and methods in ELT: A teacher's guide (*3rd ed.*)*, (pp. 79–105). John Wiley & Sons Ltd.

Stivers, T. (2019) How we manage social relationships through answers to questions: The case of interjections. *Discourse Processes*, *56*(3), 191–209. https://doi.org/10.1080/0163853X.2018.1441214

Talandis Jr, G., & Stout, M. (2015). Getting EFL students to speak: An action research approach. *ELT Journal*, *69*(1), 11–25. https://doi.org/10.1093/elt/ccu037

Thomas, G. (2013). *How to do your research project* (2nd ed.). Sage Publications.

Thompson, A. (2022). Interjections in spoken ELF interactions. In I. Walkinshaw (Eds.), *Pragmatics in English as a Lingua Franca: Findings and Developments*, (pp.147–163). Walter de Gruyter, Inc. https://doi.org/10.1515/9781501512520

Wharton, T. (2003). Interjections, language, and the "showing/saying" continuum. *Pragmatics & Cognition*, *11*(1), 39–91. https://doi.org/10.1075/pc.11.1.04wha

Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press. https://doi.org/10.1017/CBO9780511635649

Wierzbicka, A. (1992). The semantics of interjection. *Journal of Pragmatics*, *18*(2-3), 159–192. https://doi.org/10.1016/0378-2166(92)90050-L

Wilkins, D. P. (1992). Interjections as deictics. *Journal of Pragmatics*, *18*, 119–158. https://doi.org/10.1016/0378-2166(92)90049-H

İki İngilizce Ders Kitabında Yer Verilen Ünlem İfadelerinin Analizi

Öz

İngilizcede ünlem ifadelerinin öğretilmesi konusuna artan bir ilgi vardır, çünkü ünlem ifadeleri sözlü ve sözsüz iletişim arasında bir köprü görevi görerek önemli bir rol oynamaktadır. Fakat, İngilizce dili öğretiminde (ELT) ünlem ifadeleri yeterince araştırılmamış bir alan olmaya devam etmektedir. Bu çalışma, iki İngilizce ders kitabındaki (orta seviyeye tekabül eden) ünlem ifadelerinin kullanımını ve sıklığını araştırıp ve karşılaştırmıştır. Veri toplamak için Yes You Can (anadili İngilizce olmayan kişiler tarafından yazılmış ve Milli Eğitim Bakanlığı tarafından finanse edilmiştir) ve Touchstone 2 (anadili İngilizce olan kişiler tarafından yazılmış ve Cambridge University Press tarafından yayınlanmıştır) seçilmiştir. Toplamda 97 diyalog analiz edilmiştir. Bulgular, ders kitaplarında öğretilen/kullanılan ünlem ifadelerinin işlevleri ve anlamları açısından iki ders kitabı arasında bazı benzerlikler olduğunu gösterse de, entegre edilen ünlem ifadelerinin çeşitliliği ve miktarı açısından bu ders kitapları arasında bariz bir fark gözlemlenmiştir. Çalışma materyal geliştirenlere, ELT öğretmenlerine, dil öğrenenlere ve yetkililere öneriler sunmaktadır.

Anahtar kelimeler: ünlem ifadeleri, ELT ders kitapları, materyal değerlendirme, konuşma dili öğretimi

Appendix A

Semantic and Pragmatic Features of Interjections in *Yes You Can*

Functions of Interjections	Interjections	Meaning	f	%
To show admiration	Wow	"Amazing"	5	71.42
To show surprise		"Interesting"	2	28.57
To show confirmation	Yeah	"Yes"	5	83.33
To show agreement		"You are right"	1	16.66
To show hesitation	Umm	"I am not sure"	1	100
To show hesitation	Hm	"I am not sure"	2	100
To draw attention	Hey	"I want to learn smth"	2	100
To show a change in cognitive process	Oh	"I see/notice/remember"	7	43.75
To show turn-taking		No precise meaning	3	18.75
To show affection		"It is bad!"	2	12.50
To show disappointment		"I am disappointed"	1	6.25
To show pleasure		"Great!"	1	6.25
To show disagreement		"No!"	1	6.25
To show annoyance		"I am angry"	1	6.25
To show change in cognitive process	Aha	"I see"	1	100

Appendix B

Semantic and Pragmatic Features of Interjections in *Touchstone 2*

Functions of Interjections	Interjections	Meaning	f	%
To show confirmation	Yeah	"Yes"	59	48.76
To show agreement		"You are right"	38	31.40
To give feedback		"I am listening"	13	10.74
To show turn-taking		No precise meaning	5	4.13
To ask for confirmation		"Is it so?"	3	2.47
To show hesitation		"I am not sure"	2	1.65
To show self-confirmation		"I am sure"	1	0.82
To show turn-taking	Oh	No precise meaning	37	43.52
To show a change in cognitive process		"I see/notice/remember"	28	32.94
To show surprise		"Interesting"	9	10.58
To show disappointment		"I am disappointed"	6	7.05
To show affection		"I like it"	2	2.35
To show affection		"It is bad"	2	2.35
To show annoyance		"I am angry"	1	1.17
To show a thinking process	Um	"I am thinking"	18	62.06
To show hesitation		"I am not sure"	7	24.13
To show a transition to the point		"I want to say smth"	4	13.79
To show a thinking process	Uh	"I am thinking"	9	60
To show hesitation		"I am not sure"	3	20
To show a transition to the main topic		"I want to say smth"	3	20
To show a change in cognitive process	Oh yeah	"I see/notice/remember"	8	57.14
To show confirmation		"Yes"	4	28.57
To show agreement		"I agree with you"	1	7.14
To ask for confirmation		"Really?"	1	7.14
To show surprise	Huh	"Interesting"	12	85.71
To ask for confirmation		"Right?"	2	14.28
To give feedback	Uh-huh	"I am listening"	5	50
To show turn-taking		No precise meaning	5	50
To show surprise	Wow	"Interesting"	4	66.66
To show admiration		"Amazing"	2	33.33
To draw attention	Hey	"I want to learn smth"	4	80
To show annoyance		"I am angry"	1	20
To show surprise	Gosh	"Interesting"	3	75
To show disturbance		"It is disturbing"	1	25

Functions of Interjections	Interjections	Meaning	f	%
To show pleasure	Mmm	“It is delicious”	2	100
To show hesitation	Hmm	“I am not sure”	2	100
To show admiration	Oh wow	“Amazing”	1	100
To show disgust	Ugh	“It is disgusting”	1	100
To show a change in cognitive process	Ooh	“I notice”	1	100
To show anger	Oooh	“I am angry”	1	100
To show surprise	Boy	“Interesting”	1	100

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
2024, Vol. 41-3, No. 1
https://doi.org/ 10.52597/buje.1415026

Change of Epistemic Stance of Transnational Pre-Service Teachers in-and-through Collaborative Data-led Reflective Dialogues[a]

Gülşah Uyar[b] and Nilüfer Can Daşkın[c]

Abstract

This study explored collaborative data-led reflective dialogues of transnational PSTs by using Multimodal Conversation Analysis. The study draws on the screen recordings of a Virtual Exchange (VE) project (n=72), which involves collaborative online task design, implementation of tasks by L2 students, and reflection on students' performance and their experience of the VE project. The close examination of the data showed that through collaborative reflective dialogues, PSTs (i) raised awareness of their practice and generated knowledge; (ii) made a connection between theory and practice; (iii) became more aware of their epistemic stance; (iv) identified and described a problem and found solutions to these problems (Farrell, 2015). Therefore, micro-moments of learning and understanding were created in-an- through reflection and it was revealed explicitly through reference to lack of knowledge in the past. This study has implications for teacher education programs, which should encourage reflective dialogues of transnational PSTs to create opportunities for teacher learning.

Keywords: teacher education, teacher learning, data-led reflection, power asymmetry, collaboration

Article info

Received: 04.01.2024
Revised: 17.02.2024
Accepted: 27.02.2024
Published online: 22.04.2024

Introduction

With the emergence of sociocultural approach to learning and teaching, teachers and students are now seen as social agents in interaction, which has led to changes in teacher education programs (Wright, 2010). Freeman and Johnson (1998) define teacher education as a term explaining how a teacher learns to teach and how teacher trainers educate trainees to be teachers professionally by linking expert knowledge and experiential knowledge, helping teachers to gain new understandings of their teaching and themselves (Johnson & Golombek, 2016). Opfer and Pedder (2011) suggest that four orientations lead to better teacher learning: (1) field and classroom experience; (2) opportunity for reflection; (3) opportunity for understanding oneself in a secure environment under challenging or novel circumstances; (4) applied knowledge about teaching and learning. Therefore, reflection has taken its place in teacher education programs to provide pre-service teachers (PSTs henceforth) with the experience they need.

The word reflection originated from a Latin word which involves the act of looking back and noticing. "Reflective practice is: A cognitive process accompanied by a set of attitudes in which teachers systematically collect data about their practice, and while engaging in dialogue with others use the data to make informed decisions about their practice both inside and outside the classroom." (Farrell, 2015, p.140). In teacher education, reflective skills encompass the ability to find solutions to questions (Jay & Johnson, 2002) and problems observed (Dewey, 1933) drawing upon past experiences with a future orientation. Thus, it enables teachers to critically observe their practice and gain new insights. This underscores the cruciality of equipping teachers with reflective skills in teacher education (Beauchamp, 2015; Brookfield, 2017; Farrell, 2016b; Kumaravadivelu, 2006), and many scholars advocate integrating reflection into teacher education programs (e.g., Dikilitaş & Comoglu, 2022; Turhan & Kirkgöz, 2021). Reflective dialogue between trainer and trainees has been explored through Conversation Analysis (CA henceforth) (Keogh, 2010; Skovholt et al., 2019; Walsh & Mann, 2015; Waring, 2013,

[a] *This article is based on the doctoral dissertation of the corresponding author.*

[b] *Corresponding author, Erzincan University, Department of Foreign Language Education, gulsahuyr@gmail.com, ORCID: 0000-0002-5023-9971*
[c] *Hacettepe University, Department of Foreign Language Education, can_nilufer@yahoo.com, ORCID: 0000-0002-7738-0481*

To cite: Uyar, G. & Can Daşkın, N. (2024). Change of epistemic stance of transnational pre-service teachers in-and-through collaborative data-led reflective dialogues. *Boğaziçi University Journal of Education, 41-3*(1), 35-51. https://doi.org/ 10.52597/buje.1415026

2014, 2017) and other domains of research (Copland, 2008, 2011; Copland et al., 2009; Golombek, 2010; Strong & Baron, 2004). Nevertheless, there is a dearth of empirical studies which investigate how reflective practice enhances teaching (Ayoobiyan & Rashidi, 2021) and what happens when student teachers reflect on the implementation of their designs as a group without the guidance of a trainer. To wit, encouraging transnational student teachers to reflect on their task design experience and student performances without the presence of any mentor/trainer or someone who implies epistemic asymmetry may create opportunities for teacher learning. A close examination of such a practice merits further study as it may contribute to better understanding of reflective talk in LTE. We believe that without an urge to fit their ideas to a trainer's (Beck & Kosnik, 2002; Bonilla & Rivera, 2008; Farr, 2010) or without guidance and assessment of a trainer, how PSTs topicalize and generate learning merits exploration. CA enables to explore the "black box" of such dialogues (Ishino, 2018) and can provide evidence for the change of epistemic stance. With these in mind, this study sets out to be the first to explore epistemic change of PSTs in their reflective dialogues in situ. To do so, collaborative data-led reflective dialogues of transnational pre-service teachers are explored by using robust methodological tools of multimodal CA.

Literature Review

Reflective practice has gained popularity with the social turn in SLA, as a result of which teachers have come to be considered active agents in social acts rather than solely information-providers. According to Dewey (1933), reflective practice refers to "active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it" (p.6). Teacher reflection involves observing one's own actions systematically to make criticism and finding alternative action solutions to their criticism (Korthagen, 2014). Hence, reflection should not be reduced to revising past actions as it requires teachers to criticize themselves (Hickson, 2011) so that they can finetune their ideas and practices. Also, Mau and Harkness (2020) advocate for collaborative reflection instead of individual reflection. Although reflection is foregrounded by looking at back, it is supposed to involve future orientation to reach a lifelong critical level in reflection (Turhan & Kirkgöz, 2021). There exists conflict over the definition of reflection and its practice. Accordingly, Walsh and Mann (2015, p.352) problematized reflective practices: "it is (i) insufficiently data-led; (ii) heavily focused on the individual at the expense of collaborative options; (iii) dominated by written forms of reflection; (iv) lacking in detail about the nature and purposes of reflective tools."

Depending on the time it is practiced, reflective practice can be categorized as reflection-in-action, reflection-on-action, or reflection-for-action (Dewey, 1933). Reflection-in-action refers to the practice of finding a solution to a problem by the time it occurs, and reflection-on-action refers to recalling of teaching practice. Lastly, reflection-for-action refers to the mode of reflection which involves considering possible problems to be encountered in the future, and accordingly, making plans. This kind of reflection can also derive from the experience gained through reflection-in-action and reflection-on-action. (Farrell, 2015). PSTs' reflection consists of three processes: (1) description, (2) interpretation, and (3) alternatives. In the first step, PSTs describe a situation, then they evaluate the situation and lastly they come up with alternative action proposals (Kleinknecht & Gröschner, 2016). There are two levels of reflection: individual and collaborative (Murray, 2010). Individual reflection refers teachers' reflection on their own while collaborative reflection requires teachers to share their reflections with other teachers. Turhan and Kirkgöz (2021) found that collaborative reflection helped PSTs to extend their understanding affording them to examine a situation from different perspectives. Farrell (2016a) found out that novice teachers could overcome the difficulties of transition from theoretical knowledge to practice in their initial experience of teaching through a teacher reflection group. On the other hand, Farrell (2016b) avoids making the claim that reflection results in better teaching performance. but concludes that through reflection, teachers' awareness of philosophy, principles, theory, and practice was raised, which may enhance teaching.

Reflective teaching, action research, and teacher research indicated the importance of teachers' experiences and reflection, and how experiences and reflection inquiries lead to professional development. Through reflective practice, teachers' less observable practices come into view (Freeman, 2016) and it triggers professional development (Korthagen, 2014; Schön, 1983). Reflective practice enables teachers to explore their practice from the perspective of students and to advocate for development in their practices (Dikilitaş, 2015), to more extensively develop an understanding of teaching (Crandall & Christison, 2016), to increase their self-efficacy in teaching (Pedro, 2005), to improve their resilience (Ayoobiyan & Rashidi, 2021; Hong, 2012) to compare their previous knowledge and their new experience and to redesign their knowledge based on this comparison (Burton, 2009), to raise awareness of routinised practices of professionals (Caswell & Dall, 2022), to have growing identity (Dikilitaş & Mumford, 2023) and to improve their knowledge and practice (Richards,

1995). Besides, it provides opportunities for teacher learning (Balaman, 2023) and constructs teacher learning (Waring, 2017) as reflection affects not only behavior but also the cognition of teachers (Huth et al., 2019). Reflective practice aims to (1) raise individuals who do not consume knowledge but generate it, (2) bridge the gap between theory and practice, (3) make teachers more aware of their beliefs and practice through evidence, (4) identify, describe and solve a problem through evidence-based approach, (5) increase self-esteem and self-confidence of teachers and (6) raise resilience of teachers by preparing them for future problems and changes (Farrell, 2015).

Trainer/mentor-teacher reflection bears asymmetric power (Harris et al., 2019; Kim & Silver, 2021; Vasquez, 2004; Waring, 2014) by which reflective dialogue is prone to be affected by trainer/mentor guidance, questions and assessment. To cultivate reflection, teacher trainers could be less directive and give enough space for PSTs' participation (Copland et al, 2009). However, giving enough space may not diminish the asymmetry. For instance, Veen and de la Croix (2016) indicated in their study that PSTs could manage the topic, take turns and interrupt turns when they presented a case, however tutors navigated the interaction in general by opening and closing activities, facilitating reflection and leading discussion. Furthermore, Skovholt (2018) found that although the teacher in the study creates space for students to convey their ideas, her pedagogical goal may urge her to guide students to give preferred answers in feedback oriented conversation. Namely, trainers tend to manipulate trainees' ideas and put across their ideas. In other words, trainers' ideas surpass what trainees think and trainees' ideas are only valued if they resonate with the ideas of the teacher trainer (Copland et al., 2009). PSTs may not show if they do not agree with the trainers and tend to seem to agree (Beck & Kosnik, 2002; Bonilla & Rivera, 2008; Farr, 2010). The asymmetry may also lead trainers and PSTs to apply some face-saving strategies. In a recent study, Bjørndal (2020) deduced that when PSTs get critical feedback, they use various face-saving strategies as critical reflection is seen as a face-threatening act while Vasquez (2004) discovered that trainers/mentors employ various politeness strategies to save teachers' face in post observation meetings. Besides, Waring (2017) concluded that mentors tend to "go general" as an interactional strategy in post-observation conversations to depersonalize advice.

Heritage (2013, p. 370) asserts that "within Conversation Analysis (CA), research into epistemics focuses on the knowledge claims that interactants assert, contest and defend in and through turns-at-talk and sequences of interaction." The knowledge of interactants may vary from unknowing (K-) and knowing (K+) ends of gradient (Heritage, 2012a, 2012b, 2013) and epistemic status can be claimed and shown in the unfolding of conversation in situ. Sacks (1992) differentiates between claim and demonstration (of understanding) in his seminal work. While claim of understanding does not guarantee that understanding occurs or not, understanding is somehow shown in interaction in demonstration of understanding. Then, Koole (2010) extends the notions as claim and display of knowing and understanding. Lately, Sert and Walsh (2013) investigated claim of insufficient knowledge (CIK [also see Beach & Metzger, 1997 for CIK in courtrooms])in an educational context. It is sufficient to say that epistemic is not a static notion and it may change from K- to K+ in situ, which may be shown, for instance with the change of state token "oh" (Heritage, 1984).

Against this background, this study aims to reveal how the epistemic status of PSTs was revealed and changed in and through collaborative reflective dialogues without a presence of a teacher trainer. Therefore, building on the growing literature on reflection, the study explores video-mediated data-led reflective dialogues of transnational PSTs and sets out to bring conservation analytic evidence for epistemic change of PSTs.

Method

Research Context and Participants

The data of this study came from the DIGITASK4IC project. The project aims to create a sustainable application for virtual exchange tasks to improve interactional competence (IC) of second language learners. It lasted 18 months and data were collected in the academic year 2021-2022. 18 researchers from four partner universities from three countries worked for the project (Austria, Spain, and Türkiye). Depending on the courses offered, three partner university (Austria, Spain, and Türkiye) and two non-partner universities (Tunisia and Türkiye) were involved in data collection process.

In this study, there were two broad groups of participants: PSTs and students. PSTs were registered at the partner universities of DIGITASK4IC project. In the scope of the classes they were enrolled to, they were supposed to improve their skills regarding online task design. Students were enrolled to classes at different

universities than partner universities of the project which aims to better their English IC. For the purposes of this study, only the data coming from PSTs' reflection meetings were used in this study. In other words, as the secondary group participants, students who performed the tasks provided data to PSTs for reflection. It is this aspect that makes the reflective dialogues data-led. PSTs reflected on their task design and implementation by using the actual data coming from student performances rather than relying on intuitions and hypothetical practices. Besides, they did so collaboratively by engaging in meaning making in and through interaction. Therefore, L2 IC of the students was not tracked for this study.

The number of PSTs was 72 and they formed 21 groups of 3-4. There were 17 PSTs from Austrian, 13 PSTs from Spanish, and 42 from Turkish university. They were BA or MA students in language departments, and they were trained to use online tools and prepare online tasks. PSTs were required to design a virtual exchange task in groups for the students. On the other hand, students in dyads were required to perform tasks created by PSTs. The number of students was 60 and they constituted 30 dyads.

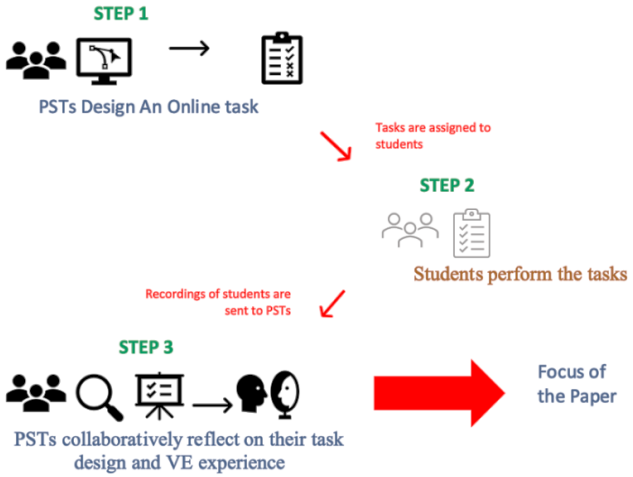
Data Collection

All the participants were required to have virtual exchange with their peers within the scope of the classes they were enrolled. The virtual exchange was conducted on MS Teams and built-in screen recording tool of the program was utilized besides licensed Screencast-O-Matic (SoM). All the participants had access to license of SoM and each participant recorded their own screen to decrease the risk of data loss.

Before data collection process started, pre-service teachers were trained on common creative (CC) licenses, designing an online task and how to use Digitask web application. A guideline for work packages were introduced to the PSTs. By the time all the preparation by the project team was completed and PSTs were prepared to design an online task, data collection process started. Data was collected in three steps: (i) Task Design Conferences, (ii) Task Engagement, (iii) Reflection (see Figure 1 below). The current study deals with the data which comes from the final step of data collection procedure, namely reflection. Nevertheless, to provide a comprehensive picture of the data collection, all steps are touched upon in the paper.

Figure 1

Data Collection Steps



Throughout the steps of data collection, the participants were instructed to record their screen by using SoM and built-in screen recording tool of MS Teams. In the first step, PSTs engaged in virtual exchange meetings to design a virtual exchange task with their group members on Digitask web application. In sum, they had five virtual exchange meetings to design a task with the teacher trainers' feedback. There was no limit for duration of the meetings and duration varied across groups. PSTs were supposed to prepare tasks according to the task design criteria. According to the criteria, they needed to prepare task on the topic assigned to their group. Topics are as follows: "Getting to know each other", "Language Learning Biographies", "Language in the Public Space", "Language Varieties", "Lingua Franca", "Communication styles", "Multilingualism", "Inclusive Language" and "Final Reflection Task". The tasks should have real-life learning objectives with clear goals, create learning opportunities for students, integrate different skills, be suitable for online task settings, and last about 30 minutes.

Once the PSTs finalized their task design at the end of five virtual exchange meetings, their tasks were assigned to the students. Each task was assigned to eight or nine dyads of students. In this way, PSTs could see how their task was performed by various student dyads.

Preceding the third step of data collection, which is reflection, screen-recordings of the students were sent to PSTs to furnish data for their reflection. PSTs watched the screen-recordings individually, and lastly, they had two meetings to reflect on their task design experience, collaboration, and student performances (see Figure 1 above). Reflections were conducted without a teacher trainer or mentor. However, guiding questions were presented to PSTs to enable them to reflect. In the first reflection meeting, PSTs were supposed to share their observations on the recordings of student performances and in the second reflection meeting, they assessed the entire experience referring to guiding questions. Thus, the participants needed to reflect on collaboratively with their team members and they were supposed to so by referring to students' actual performance data, thereby making it data-led. Therefore, the setting urged the participants for collaborative data-led reflection. The work package for each reflection meeting and guiding questions on the guideline are as follows:

Table 1

Guide for Reflection

Meeting	Guiding questions
Meeting 6	Share your observations on the recordings you viewed (How was the task implemented? / Did learners behave like you expected? Did learners do something you had not anticipated? / Was the task successful?)
Meeting 7	Group Reflection Part 2 & Closing. Collaboratively assess the entire partnership this semester with reference to the lessons learned (What did you learn? What would you do differently in a future collaborative task design? Was the project useful for your future professional life as a language teacher?)

Data Analysis

The data from the reflection meetings is approximately 18 hours. They study adopts Multimodal CA to analyze the data. Multimodal CA is a bottom-up research methodology that recognizes that interaction is a complex interplay of verbal utterances, prosody, gestures, body language, gaze and artefacts. Multimodal CA approaches the data without any pre-determined hypothesis or questions (Hutchby & Wooffitt, 2008). Therefore, unmotivated looking and transcription and lastly line-by-line rigorous analysis underlie the steps of multimodal CA. Accordingly, first, orthographic transcriptions were prepared. In this step, only verbal utterances are transcribed and there is no detail of prosody and embodiment. Then, through unmotivated looking, that is letting the data speak without imposing pre-determined categories and external theories (ten Have, 2007), commonly occurring practices were identified. Next, collection of these practices was built and lastly, detailed transcripts were prepared following the transcript conventions (Jefferson, 2004; Mondada, 2018). We have found out that PSTs claim lack of knowledge referring to their epistemic stance in the past (see Table 2 below). Unknowing status in the past was topicalized with various lexicon and syntactic markers such as "didn't know" (7 cases), "I wasn't thinking/I didn't think" (7 cases), "had no idea (2 cases), "didn't expect" (1 case), "didn't foresee" (1 case), "never grasped" (1 case), "imagined" (1 case), "wasn't aware of" (1 case), and "was abstract at first" (1 case) . Four representative extracts from the dataset are presented to show epistemic change of PSTs.

Table 2

Collection- Lack of Knowledge in the Past

	Number of Cases
Didn't know	7
I wasn't thinking/I didn't think	7
Had no idea	2
Didn't expect	1
Didn't foresee	1
We never grasped	1

	Number of Cases
Wasn't aware of	1
Abstract at first	1

Findings

We selected four representative extracts from our dataset to show how micro-moments of learning and understanding were created in reflective dialogues of PSTs. Representative extracts were chosen regarding various lexicon to refer to lack of knowledge in the past and topic of teacher learning. By doing so, we aim to show the range of teacher learning across topics, as well. We have found that by reflecting on a gap or problem in the past, PSTs display their awareness of knowledge gaps in the past. Therefore, we conclude that they have extended their learning trajectories as their epistemic status has changed from K- (less knowledgeable) to K+ (more knowledgeable). In the transcripts, we have followed Jefferson (2004) and Mondada (2018) transcript conventions (see Appendix).

In extract 1, PSTs claim learning or extended knowledge of Common Creative licenses on which they were trained by teacher trainers before data collection procedure started. They use unknowing epistemic markers (K-) to show their lack of knowledge or topicalize how their knowledge is extended.

Figure 2

Extract: Didn't Know

```
1  LIN: yeah (0.3) i learned that too (.) now about (.)licenses
2       and everything ♣i mean i already knew♣ (0.8) um
   der                 ♣---------nods-----♣
3       (1.0)
4  DER: we didn't [know
5  LIN:           [knew about that but not in detail and
6       now ♣i'm (1.0) also more conscious of that [aspect♣
   der       ♣-----------------nods---------- ---------♣
7  EMI:                                          [yeah
```

Note. ♣: Multimodal actions of DER

The extract starts with LIN's use of an acknowledgement token (`yeah`) (Jefferson, 1984). In the same turn, she claims learning (`i learned that`) of common creative licenses. DER shows alignment with this claim of learning in line 2 with a nod. Then in line 2, LIN highlights her knowing epistemic status (K+) (Heritage, 2013) in the past (`i already knew`). After a lengthy silence of 1.0 second, DER takes the turn and points to their insufficient knowledge in the past in line 4 (`we didn't [know`). She uses the first-person plural pronoun to involve her peer EMI, who is a student at the same institution as herself. DER's statement was overlapped with LIN's extension of her previous turn. In line 5, LIN highlights her epistemic status in the past as knowing (K+) once more. However, in the remaining of the turn, she indicates her knowledge of licenses was not detailed and she is more aware now which is accompanied with DER's nodding. In the last line, EMI shows alignment with an acknowledgement token.

Extract 1 shows participants' orientation to their epistemic status in the past and at the time of reflection. LIN opens the turn with a claim of learning which was then explained with comparison to her epistemic status in the past. Although LIN mentions her epistemic status as K+ in line 2 and 5, she mentions that her knowledge on common creative licenses has been extended and detailed in line 5 and 6. Her structure as "knew about that but not in detail" implies an extension in her epistemic stance. Besides, her expression of her epistemic stance from "I learned" to "I already knew" and then to "knew about that but not in detail" shows how she regulates her ideas (see Heritage, 2012a, 2012b, 2013) in and through reflection. On the other side, DER involves her peer from the same institution by using a first-person plural subject pronoun (we) and shows their lack of knowledge in the past with a negation in line 4. Therefore, it is obvious that there is a change of epistemic status of DER and maybe EMI who shows alignment with a minimal token in line 7. Therefore, this extract showed that, the process helped pre-service teachers to extend their already existing knowledge and to form new knowledge. All in all, they had

the chance to evaluate their experience and regulate their epistemic status during reflective dialogues with their peers.

In their previous dialogues before extract 2, PSTs have problematized the disappearance of their instructions on the interface of the application and how this affected students' performance. Before the extract starts, GOK reads one of the guiding questions given to them: what would you do differently in a future collaborative task design? The extract starts with his alternative action proposal on the materials they have chosen.

Figure 3

Extract 2: We didn't Foresee

```
1  GOK: ♣↑one♣ thing (.) for sure (0.8) i would prepare
        ♣points♣ with index finger
2       my materials (0.7) that (0.7) can be (0.4)
3       understood (.) by their own
4       (1.5)
5       ♥if ♠instructions are lost (.) the students need to
  jas   ♥nods--->
  sev       ♠nods-->
6      ♣understand the task♠ (.) as (1.0) ↑well as they can(.)
       ♣shakes head side to side--->
  sev                    --->♠
7      by just looking at the materials (0.8) so this could be
8      a solution♥ for our case♣(1.0) because er in ↑our table
  jas         --->♥
                              --->♣
9       (.) they (0.9) discovered what to do (1.3) ♣ they see♣
                                                  ♣---1----♣
       1: points to the top of the frame
10      some gender err exclusive words and they ♣ (0.8) er
                                                 ♣draws a round
                                                 with index finger-->
11     ♣found (.) the opposite one ♥they found the gender♥
      -->♣
  jas                                ♥---------nods----------♥
12     neutral one ↑it was clear (0.3) er maybe (.) our (0.5)
13     ♣other♣ materials (0.6) could have been (.) as clear
       ♣shakes♣ head to right
14     (0.3) as ♥ (.) the last ones♥ (1.2) this could can be
  jas           ♥--------nods-------♥
15     a change
16     (4.0)
17     ♥because we didn't foresee♥ (.)this problem (.)to happen
  jas  ♥--------------nods----------♥
18     (.) but it ↑happened ♥ (0.8) so♥ (1.0) maybe i err
  jas                       ♥---nods---♥
19     think.hh (.) wider
```

Note: ♣: Multimodal actions of GOK, ♥: Multimodal actions of JAS, ♠: Multimodal actions of SEV

Extract 2 starts with GOK's highlighted initiation for alternative action proposal (`i would prepare`) (Kleinknecht & Gröschner, 2016; Korthagen, 2010). In lines 2 and 3, he suggests choosing self-explaining materials for the task. YAS and SEV show alignment by nodding in line 5. After a lengthy silence in line 4, he starts elaborating on his proposal in line 5. From line 5 to 7, he elaborates on proposing self-explaining materials for the task with an if clause. In line 7 and 8, he declares his alternative action proposal as a solution to the problem (`this could be a solution`). Then in the remaining of line 8, he starts giving an account of his alternative action proposal. From line 8 to 12, he describes a situation he observed in students' performance. He gives account by referring to how students managed to discover what they were required to do in line 9 (`they (0.9) discovered what to do`) and in line 11 (`found (.) the opposite one`) and YAS shows alignment

by nodding. Then in line 12, he adopts an evaluative stance and shares his ideas about the table (it was clear). After a mitigation token (maybe) in line 12, he repeats his alternative action proposal in line 13 and 14. His alternative action proposal was one more time aligned by YAS in line 14 by nodding. In the remaining of the line 14 and in line 15, GOK highlights his alternative action proposal as a solution to the problem they have encountered once more. After a quite long pause in line 16, GOK gives an account (because we didn't foresee (.)this problem (.)to happen) and verbalizes their K- (Heritage, 2012a, 2013) epistemic stance in the past. Then, in line 18, he declares that they have encountered the problem which they did not foresee (but it ↑happened). Following a transition marker (so) and mitigation (maybe), he refers to his extended epistemic stance in the present (i err think.hh (.) wider) in lines 18 and 19.

The extract shows that their observation of students' performance and reflection helped pre-service teachers to extend their perspectives of teaching. As already mentioned by GOK, in the last line of the extract, all the process and critical reflection on it enabled them to have a wider point of view as they have encountered unexpected situations in students' implementation. So, critical reflection played an important role in their knowledge formation and extension.

In extract 3, PSTs refer to the problem they have encountered during their virtual exchange experience regarding a time difference between countries of partner universities and reflect on this problem with a future orientation.

Figure 4

Extract 3: Wasn't Thinking

```
1   VIK: uh >especially talking about< ca international collaboration
2        (0.5) that some things aren't that obvious so
3        ♣if you're talking about okay we should
  mer:   ♣plays with his chin ---> line 10
4        meet ♦we (.) organize a meeting (.) at the first time
  eli:        ♦smiles--->line 15
5        i wasn't even thinking about a time difference
6   ELI: $hah hah$
7   VIK: it was like okay yeah that's true we have to organize er um
8        (.)on that regard as well (0.6) so all the orga-
9        organizational aspects (.) of (0.5) an international
10       collaboration↓ (.) and i think♣
  mer:                              --->♣
11       this is quite helpful if you think about maybe doing some (.)
12       things (.) similar with your future students (0.7) because
13       ♥there's so much around (.)the task itself♥
         ♥----------rolls her hands----------------♥
14       (0.6)that ♣ (.) shouldn't be forgotten (0.4) and i think♣
  mer:             ♣--scratches his head with left index finger--♣
15       it's good to experience that er one ♥time♦ (1.0)♥♦
                                            ♥nods and smiles♥
  eli:                                          ♦--nods--♦
  eli:                                               --->♦
```

Note: ♣: Multimodal actions of MER, ♥: Multimodal actions of VIK, ♦: Multimodal actions of ELI

Extract 3 starts with an elaboration of VIK on international collaboration. In line 2, she touches upon the unclarity of the process (some things aren't that obvious) by adopting an evaluative stance. With reference to the possibility of organizing an international meeting, she verbalizes her K- epistemic status in the past about considering the time difference between various countries (i wasn't even thinking about a time difference) in line 5 and by doing so she refers to a gap and claims lack of experience. Then, in line 6, ELI responses to VIK's claim with a laughter. From line 7 to 10, VIK refers to their own experience and explains what they needed to do as a part of international collaboration. In 11 and 12, we see how VIK adopts an evaluative stance (this is quite helpful) towards the experience and evaluates it to be helpful for their future career (maybe doing some (.)things (.) similar with your future students). What follows is her statement about what she has gained awareness of and her action proposal (♥there's so much around (.)the task itself♥ that ♣ (.) shouldn't be forgotten). Lastly, in line 15,

she adopts an evaluative stance one more time (it's good to experience that er one ♥time♦) and clearly states that to experience an international collaboration was good (also see Lortie, 1975).

The extract shows that although VIK was not aware of considering the time difference in an online collaboration at the beginning, her awareness was raised after her experience as a pre-service teacher in the project. As a prospective teacher, she orients to the future and displays her takeaway of considering time difference in the case of conducting such a process in her future career. We witness an epistemic change from K- in the past to K+ in the present as she refers to a gap in her epistemic stance in the past. So, with the process of task design, feedback, students' data and reflection, VIK's awareness of different aspects to be considered while conducting an online collaboration has been raised and the takeaway lesson was negotiated and became apparent during reflection.

Before Extract 4 starts, DER talks about the individual differences she has observed while watching students' recordings and how students comprehended and performed the task in unexpected ways.

Figure 5

Extract 4: Didn't Think- We Know

```
1   DER: of course we don't think that everybody thinks it's the
2        same way↑ but like ♣seeing that especially♣
                           ♣unites finger while palm♣is up
3        ♣oh i didn't think of this i didn't think of that♣
         ♣-----------puts finger on the head--------------♣
4        [((inaudible))
5   LIN: [((inaudible))
6        the small things ♥also (0.6) th the timing ♣we talked
                          ♥looks up--->
    der                                                ♣leans forward--->
7        about♣ (.) because♥ (0.4) yeah ♥we we know it with our
    der   --->♣
                       --->♥              ♥scrolls right hand--->
8        screen castomatic right now we have technical problems
9        ♣ (.)♣ it is (0.8)it can happen all the time and you have
    der  ♣nods♣
10       to♥ calculate tha:t (.) maybe ♣because (0.4) wh♣ what we
        -->♥
    der                                  ♣--slightly nods--♣
11       saw in the tasks (0.5) ♣er was also (0.4) that♣ maybe
    der                             ♣-----leans forward------♣
12       they needed more time and ♥they couldn't do: ♣ (0.4) one♥
                                    ♥-----shakes head side to side---♥
    der                                                 ♣nods--->
13       situation♣ or something like that (0.7) ♣because of♣ te
    der       --->♣                                 ♣slightly nods♣
14       (.) technical issues ♣ (0.5) and that's also♣ something i
    der                          ♣strongly nods and averts gaze♣
15       (.) i $learned out of this$ haha (.)♣because if you are
    der                                          ♣smiles --->
16       in a setting♣ especially ♥(0.7) erm in a setting (0.3)
    der           --->♣
                                    ♥looks up --->
17       that the students are not ↑used to♥ ♣ (.) because they
                                         --->♥
    der                                      ♣strongly nods--->
18       have to get to know each other (0.4) and (.) and♣ stuff
    der                                                  --->♣
19      it's (1.0) ♣it's something tha:t (.) that is important
    der:             ♣leans forward--->
20      to think about♣ (.) what (.) how much time do they nee:d
    der:            --->♣
```

Note: ♣: Multimodal actions of DER, ♥: Multimodal actions of LIN

The extract starts with DER's display of their epistemic stance (of course we don't think) (also see Koole, 2010). She mentions that they are aware of the fact that everyone thinks in a different way. However, she also highlights how they have observed this fact when watching student performances (seeing that especially). In line 3, she refers to her K- epistemic status in the past (oh i didn't think of this i didn't think of that), which she became aware of after watching students' unexpected ways of comprehension and behavior. In line 6, LIN initiates by adding up to DER's problematization and starts delivering a proposal (the small things ♥also (0.6) th the timing) and giving account (because) in line 7. Then, she indicates their (+) epistemic stance (we know it) at the time of speaking by referring to a technical problem with screen-recording software (screencast-o-matic) in their own virtual exchange experience (our screen castomatic right now we have technical problems). In line 9, she verbalizes that the problem they have encountered may always occur (it can happen all the time). Next, she proposes that they need to consider this (you have to♥ calculate tha:t) in lines 9 and 10 as an action proposal. In the remaining of line 10, she initiates giving an account of her proposal in lines 10 and 11 (because (0.4) wh♣ what we saw in the tasks) by referring to students' video-recordings. In line 12, she continues account giving and points to a gap that students needed more time to perform the task (they needed more time and ♥they couldn't do:). In lines 13 and 14, she starts giving an account of the students' performance by referring to technical issues. Then, she explicitly claims learning (that's also♣ something i (.) i $learned out of this$) (Sacks, 1992) in lines 14 and 15 by referring to what she described in her turn previously. She goes on with account-giving in line 15 by elaborating on her account and adopting an evaluative stance (it's something tha:t (.) that is important) and lastly, she refers to the take-away lesson again.

We have seen in Extract 4 that although PSTs expected students to be different from each other, their perception of how unexpected they can be is extended through watching students' performance. In addition, PSTs display their experience on technical problems and the gap they have noticed while students perform the tasks, and they reflect on them. They come to the conclusion about considering the setting while defining time to perform a task. The extract showed how PSTs orient to their own experience as a learning opportunity. How their virtual exchange experience and watching students' performance helped them to notice the gap were topicalized and learning was claimed regarding task time. Therefore, their epistemic stance has changed from K- to K+ in and through reflection.

Discussion

In this study, we have examined how data-led reflection was achieved by transnational PSTs using the micro lenses of multimodal CA. We have found that reflection creates micro-moments of learning and understanding from the participants' own orientation.

As proposed by Kleinknecht and Gröschner (2016), reflective dialogues of PSTs involve three processes: description, interpretation, and alternatives. PSTs first describe a problem or a gap, adopt an evaluative stance and find alternative solutions to problems. In doing so, they deliver accounts, claim lack of knowledge in the past, and K+ epistemic stance in the present. To show the change in their epistemic stance, they deliberately use syntactic markers and lexicon such as "didn't know", "wasn't thinking", "didn't foresee", "didn't think", "learned" and "we know".

Collaborative reflective dialogues of the PSTs had various affordances which is in line with previous research (Farrell, 2015). First, PSTs enhanced understanding of themselves and produced new insights. For instance, in extract 2, GOK generated knowledge on task materials. By referring to other materials they had used, he proposed all the materials could be as comprehensible as possible even without instructions. Also, in extract 3, it is obvious that VIK generated insight based upon her experience regarding organizing international collaboration. The change in her epistemic stance from K- end to K+ end of the gradient enables her to generate knowledge about organizing international collaboration with her future students. Similarly, in extract 4, PSTs speak of a common problem faced by both of them and students. On the grounds of their experience with SoM (see Methodology section), they concluded to set time in an online task design by considering the setting and context as online settings might bring out technical problems. Second, they established connection between theory and practice. In the second extract, GOK drew on students' performance in his extended turn (i.e. line 9 (they (0.9) discovered what to do)) and with regard to his observation, he regulated his theoretical knowledge. Furthermore, in extract 4, although PSTs were trained on online task design, they have generated

knowledge and made connection with theory and practice. Only after they watched how their task was performed, they could realize that they needed to consider the setting and problems that students might face with. Third, they raised awareness of their knowledge and epistemic stance. One of the most explicit examples is seen in Extract 1. Following the claim of learning, LIN claims knowing in the past as well (yeah (0.3) i learned that too (.) now about (.)licenses and everything ♣i mean i already knew♣ (0.8)). In the remaining of the extract, she deploys how her knowledge was extended. Therefore, the extract showed how the participants became aware of their knowledge status in situ during their reflective dialogues. Moreover, in Extract 2, GOK claimed K- epistemic stance in the past (we didn't foresee♥ (.)this problem (.)to happen) which was followed by his explicit claim of extended epistemic status (i err think.hh (.) wider). Extract 3 indicated that VIK gained new insight about international collaboration which is obvious through her explicit claim of K- status in the past (i wasn't even thinking about a time difference). Furthermore, PSTs displayed their present and past epistemic stance regarding individual differences. They depicted that their epistemic stance was extended which is specifically shown with a surprise token (oh) (Goffman, 1978). In addition, they claimed knowing that technical problems are prone to happen in online settings (we know it with our screen castomatic right now we have technical problems). Finally, they identified problems and generated solutions to these problems. Prior to extract 2, PSTs had problematized disappearance of the instructions in their task. In extract 2, GOK proposes alternative solution to this problem by selecting more self-explanatory materials. In extract 3, VIK touched upon an organizational problem they had encountered. Then, by adopting a future-orientation, she proposed to consider time difference between countries in any international collaboration. Therefore, the problem they had experienced paved the way to teacher learning. PSTs in extract 4 problematized the time they had given to students as students could not finish the task properly. They elaborate on the problematization by giving example from their virtual exchange experience which is followed by a take-away lesson regarding timing and online settings. Thus, their identification of problem ended up with proposing solutions (Korthagen, 2014) by giving reasons.

Besides, as mentioned above, PSTs explicitly referred to students' performances in their collaborative reflection. Extract 2 and 4 displayed how their observation shaped their learning. In both extracts, PSTs' problematization of students' performance led them to generate solutions; thus, marking the crucial role of data in shaping learning (Walsh & Mann, 2015). In other saying, video has been a useful tool for observation (see Körkkö, 2019; Körkkö et al., 2019). The current study also revealed that PSTs generated solutions and knowledge on various topics based upon their observation and problematization. The range of topics varies. Extract 2 indicated that PSTs' observation and problematization of disappearance of instruction paved the way to generate knowledge on material selection. They proposed to choose more self-explanatory materials in such a setting. Extract 3 showed that PSTs generated knowledge on organizing online collaboration. They proposed to consider time difference between countries when organizing online collaboration with their future students. In extract 4, they concluded that online settings are prone to technical problems, and they need to consider this while setting time. Therefore, we see proposing alternative solutions as generating knowledge practice of PSTs as well. They topicalize learning and orient to the future despite the lack of a teacher trainer and guidance on-site. Besides, beyond the guiding questions on the guideline, they formulate their turns to give an account of the problems and alternative solutions, adopt evaluative stance, and propose alternative solutions. Namely, they do not perceive reflection as a repair tool (Hickson, 2011), but take it as a tool for teacher learning for their future career (see Turhan & Kirkgöz, 2021). Moreover, they explicitly express their future-orientedness.

In contrast to prior studies that examine reflection with a presence of a teacher trainer (e.g., Skovholt et al., 2019; Waring, 2017), it is clear that PSTs' turns are quite extended as they elaborate on their answers in the extracts without a power asymmetry (Harris et al., 2019; Kim & Silver, 2021; Waring, 2014). Thus, we propose that PSTs could express themselves without an urge to fit their ideas to a teacher trainer or they found their own way without any manipulation by a professional (Beck & Kosnik, 2002; Bjørndal, 2020; Bonilla & Rivera, 2008; Copland et al., 2009; Farr, 2010; Skovholt, 2018; Vasquez, 2004; Veen & de la Croix, 2016; Waring, 2017) which led to different turn structure unlike IRF (Initiation, Response, Feedback) patterns in a typical teacher directed settings (Mehan, 1979).

The findings of the study indicated that the epistemic status of the participants moved from the unknowing end to the knowing end of the gradient (Heritage, 2012a, 2012b, 2013) upon various teaching practices, which was claimed or shown though talk-in-interaction (Koole, 2010; Sacks, 1992). All in all, reflection led to a claim of teacher learning (Balaman, 2023; Opfer & Pedder, 2011) and the findings support the proposal of integrating reflection into language teacher education programs (e.g., Dikilitaş & Comoglu, 2022; Turhan & Kirkgöz, 2023). As detailed above, through collaborative reflective dialogues, PSTs (i) raised awareness of

themselves and generated knowledge; (ii) made connection between theory and practice after watching students' performance; (iii) became more aware of their knowledge and epistemic stance; (iv) identified and described a problem and found solutions to these problems (Farrell, 2015). Therefore, microanalytic lenses of CA offered insights into how PSTs reflected (also see, Ishino, 2018) and how micro moments of learning and understanding were created in and through reflection, and the study contributed to the literature on epistemics (Heritage, 2012a, 2012b, 2013; Koole, 2010; Sacks, 1992; Sert & Walsh, 2013) through a claim of insufficient knowledge in the past regarding teaching practices.

The current study presented the affordances of data-led collaborative dialogues among PSTs. It is evident that data (from student performances) and collaboration played important roles in reflective dialogues of transnational PSTs by giving them the chance to collaboratively reflect on real practices, leading to genuine and interactional reflections. PSTs could enhance teacher learning upon various topics. Furthermore, being part of a transnational group enabled them to collaborate with their peers from diverse cultures and backgrounds, thereby globalizing teacher education (Koskela et al., 2023). We conclude that data-led collaborative reflection (Farrell, 2016a, 2016b; Turhan & Kirkgöz, 2021; Walsh & Mann, 2015) should be encouraged as an integral part of LTE programs rather than merely perceived as a repair tool (Farrell, 2018). We also put forward that various practices through which PSTs can gain understanding on educational ecology should be integrated to LTE programs and PSTs should be urged to reflect on them to generate knowledge and be aware of their epistemic stance. Therefore, for further research, we recognize the value in exploring reflective dialogues on various contexts more thoroughly by adopting conversation analytic research methodology. Future studies can also focus on benefits of data-led, collaborative, and transnational reflection and how teacher learning is shaped in situ.

Code of Ethics

Ethical Clearance was granted from the Ethical Committee at Hacettepe University on 12.04.2022 with the number of decision E-35853172-300-00002138909.

Acknowledgements

We would like to express our gratitude to Dr. Ufuk Balaman for his supervision, support and guidance on the dissertation to be submitted by the first author and for leading the Digitask project, which provided the setting for this research. We are also grateful to Dr. Nur Yiğitoğlu Aptoula and external reviewers for their invaluable contributions to this article.

References

Ayoobiyan, H., & Rashidi, N. (2021). Can reflective teaching promote resilience among Iranian EFL teachers? A mixed-method design. *Reflective Practice*, *22*(3), 293–305. https://doi.org/10.1080/14623943.2021.1873758

Balaman, U. (2023). *Conversation analytic language teacher education in digital spaces*. Springer Nature.

Beach, W. A., & Metzger, T. R. (1997). Claiming insufficient knowledge. *Human Communication Research*, *23*(4), 562–588. https://doi.org/10.1111/j.1468-2958.1997.tb00410.x

Beauchamp, C. (2015). Reflection in teacher education: Issues emerging from a review of current literature. *Reflective Practice*, *16*(1), 123–141. https://doi.org/10.1080/14623943.2014.982525

Beck, C., & Kosnik, C. (2002). Components of a good practicum placement: Student teacher perceptions. *Teacher Education Quarterly*, *29*(2), 81–98.

Bjørndal, C. R. P. (2020). Student teachers' responses to critical mentor feedback: A study of face-saving strategies in teaching placements. *Teaching and Teacher Education*, *91*, 103047. https://doi.org/10.1016/j.tate.2020.103047

Bonilla, S., & Rivera, P. M. (2008). Mentoring in pre-service teaching: From reflection on practice to a didactic proposal. *Actualidades Pedagógicas*, *1*(52), 79–90.

Brookfield, S. D. (2017). *Becoming a critically reflective Teacher*. John Wiley & Sons.

Burton, J. (2009). Reflective Practice. In A. Burns & J. C. Richards (Eds.), *The Cambridge Guide to Second Language Teacher Education* (pp. 298–307). Cambridge University Press.

Caswell, D., & Dall, T. (2022). Using conversation analysis to develop reflective practice in social work. *Qualitative Social Work*, *21*(6), 1290–1307. https://doi.org/10.1177/14733250221124210

Copland, F. (2008). Deconstructing the discourse: Understanding the feedback event. In S. Garton, & K. Richards (Eds.), *Professional encounters in TESOL*. Palgrave Macmillan. https://doi.org/10.1057/9780230594173_1

Copland, F. (2011). Legitimate talk in feedback conferences. *Applied Linguistics*, *33*(1), 1–20. https://doi.org/10.1093/applin/amr040

Copland, F., Ma, G., & Mann, S. (2009). Reflecting in and on post-observation feedback in initial teacher training on certificate courses. *English Language Teacher Education and Development*, *12*. 14–23.

Crandall, J. A. & Christison M. (Eds.). (2016). *Teacher education and professional development in TESOL: Global perspectives*. Routledge.

Dewey, J. (1933). *How we think*. Courier Corporation.

Dikilitaş, K. (2015). Professional development through teacher-research. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 47–55). IATEFL Research Special Interest Group.

Dikilitaş, K., & Comoglu, I. (2022). Pre-service English teachers' reflective engagement with stories of exploratory action research. *European Journal of Teacher Education*, *45*(1), 26–42. https://doi.org/10.1080/02619768.2020.1795123

Dikilitaş, K., & Mumford, S. (2023). Identity reconstruction through reflection and reflexivity: A new journey beyond the Ph.D. dissertation. *Reflective Practice*, *24*(3), 265–278. https://doi.org/10.1080/14623943.2023.2170342

Farr, F. (2010). *The discourse of teaching practice feedback: A corpus-based investigation of spoken and written modes*. Routledge. https://doi.org/10.4324/9780203846742

Farrell, T. S. C. (2015). *Promoting teacher reflection in second language education: A framework for TESOL professionals*. Routledge.

Farrell, T. S. C. (2016b). Anniversary article: The practices of encouraging TESOL teachers to engage in reflective practice: An appraisal of recent research contributions. *Language Teaching Research*, *20*(2), 223–247.

Farrell, T. S. C. (2016a). Surviving the transition shock in the first year of teaching through reflective practice. *System*, *61*, 12–19. https://doi.org/10.1016/j.system.2016.07.005

Farrell, T. S. C. (2018). Operationalizing reflective practice in second language teacher education. *Journal of Second Language Teacher Education*, *1*. 71–88.

Freeman, D. (2016). *Educating Second Language Teachers*. Oxford University Press.

Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge-base of language teacher education. *TESOL Quarterly*, *32*(3), 397–417. https://doi.org/10.2307/3588114

Goffman, E. (1978). Response Cries. *Language*, *54*(4), 787–815. https://doi.org/10.2307/413235

Golombek, P. R. (2010). Dynamic assessment in teacher education: Using dialogic video protocols to intervene in teacher thinking and activity. In P. R. Golombek, & K. E. Johnson (Eds), *Research on Second Language Teacher Education* (pp. 135–149). Routledge.

Harris, J., Theobald, M., & Keogh, J. (2019). Combining analytical tools to inform practice in school-based professional experience. *Journal of Pragmatics*, *143*, 255–266. https://doi.org/10.1016/j.pragma.2018.04.002

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M., Atkinson (Ed.) *Structures of Social Action*, (pp. 299–345). Cambridge University Press.

Heritage, J. (2012a). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction*, *45*(1), 1–29. https://doi.org/10.1080/08351813.2012.646684

Heritage, J. (2012b). Epistemics in conversation. In *The Handbook of Conversation Analysis* (pp. 370–394). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118325001.ch18

Heritage, J. (2013). Epistemics in conversation. In J. Sidnell, & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 370–394). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118325001.ch18

Hickson, H. (2011). Critical reflection: Reflecting on learning to be reflective. *Reflective Practice*, *12*(6), 829–839. https://doi.org/10.1080/14623943.2011.616687

Hong, J. Y. (2012). Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teachers and Teaching*, *18*(4), 417–440. https://doi.org/10.1080/13540602.2012.696044

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis*. Polity.

Huth, T., Betz, E., & Taleghani-Nikazm, C. (2019). Rethinking language teacher training: Steps for making talk-in-interaction research accessible to practitioners. *Classroom Discourse*, *10*(1), 99–122.

Ishino, M. (2018). Micro-longitudinal conversation analysis in examining co-teachers' reflection-in-action. *System*, *78*. https://doi.org/10.1016/j.system.2018.07.013

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education*, *18*(1), 73–85. https://doi.org/10.1016/S0742-051X(01)00051-8

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "Yeah"; and "Mm Hm"; *Paper in Linguistics*, *17*(2), 197–216. https://doi.org/10.1080/08351818409389201

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Pragmatics and Beyond New Series*, *125*, 13–34.

Johnson, K. E., & Golombek, P. R. (2016). *Mindful L2 teacher education: A sociocultural perspective on cultivating teachers' professional development*. Routledge. https://doi.org/10.4324/9781315641447

Keogh, J. (2010). (In)forming formal evaluation: Analysis of a practicum mentoring conversation. *Journal of Applied Linguistics and Professional Practice*, *7*, 51–73. https://doi.org/10.1558/japl.v7i1.51

Kim, Y., & Silver, R. E. (2021). "What do you think about this?": Differing role enactment in post-observation conversation. In S. Kunitz, N. Markee, & O. Sert (Eds.), *Classroom-based conversation analytic research: Theoretical and applied perspectives on pedagogy*. Springer International Publishing. https://doi.org/10.1007/978-3-030-52193-6

Kleinknecht, M., & Gröschner, A. (2016). Fostering preservice teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. *Teaching and Teacher Education*, *59*, 45–56. https://doi.org/10.1016/j.tate.2016.05.020

Koole, T. (2010). Displays of epistemic access: Student responses to teacher explanations. *Research on Language and Social Interaction*, *43*(2), 183–209. https://doi.org/10.1080/08351811003737846

Körkkö, M. (2019). Towards meaningful reflection and a holistic approach: Creating a reflection framework in teacher education. *Scandinavian Journal of Educational Research*, *65*(2), 258–275. https://doi.org/10.1080/00313831.2019.1676306

Körkkö, M., Morales Rios, S., & Kyrö-Ämmälä, O. (2019). Using a video app as a tool for reflective practice. *Educational Research*, *61*(1), 22–37. https://doi.org/10.1080/00131881.2018.1562954

Korthagen, F. A. J. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*, *26*(1), 98–106. https://doi.org/10.1016/j.tate.2009.05.001

Korthagen, F. A. J. (2014). Promoting core reflection in teacher education: Deepening professional growth. In L. Orland-Barak & C. J. Craig (Eds.), *International Teacher Education: Promising Pedagogies (Part A)* (Vol. 22, pp. 73–89). Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-368720140000022007

Koskela, T., Granö, P., & Somerkoski, B. (2023). Student teachers' experiences of practical training in a foreign culture. *Reflective Practice*, *24*(1), 71–84. https://doi.org/10.1080/14623943.2022.2139233

Kumaravadivelu, B. (2006). *Understanding language teaching: From method to postmethod*. Routledge.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. University of Chicago Press.

Mau, S. T., & Harkness, S. S. (2020). The role of teacher educators and university supervisors to help student teachers reflect: From monological reflection toward dialogical conversation. *Reflective Practice*, *21*(2), 171–182. https://doi.org/10.1080/14623943.2020.1716710

Mehan, H. (1979). 'What time is it, Denise?": Asking known information questions in classroom discourse. *Theory into Practice*, *18*(4), 285–294. https://doi.org/10.1080/00405847909542846

Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, *51*(1), 85–106. https://doi.org/10.1080/08351813.2018.1413878

Murray, A. (2010). Empowering teachers through professional development. *English Teaching Forum*, *48*(1), 2–11.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3). https://journals.sagepub.com/doi/abs/10.3102/0034654311413609

Pedro, J. Y. (2005). Reflection in teacher education: Exploring pre-service teachers' meanings of reflective practice. *Reflective Practice*, *6*(1), 49–66. https://doi.org/10.1080/1462394042000326860

Richards, J. C. (1995). Towards reflective teaching. *English Teacher Journal*, 59–63.

Sacks, H. (1992). *Lectures on conversation: Volumes I & II*. Blackwell.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books. https://doi.org/10.4324/9781315237473

Sert, O., & Walsh, S. (2013). The interactional management of claims of insufficient knowledge in English language classrooms. *Language and Education*, *27*. 542–565. https://doi.org/10.1080/09500782.2012.739174

Skovholt, K. (2018). Anatomy of a teacher–student feedback encounter. *Teaching and Teacher Education*, *69*, 142–153. https://doi.org/10.1016/j.tate.2017.09.012

Skovholt, K., Nordenström, E., & Stokoe, E. (2019). Evaluative conduct in teacher–student supervision: When students assess their own performance. *Linguistics and Education*, *50*, 46–55. https://doi.org/10.1016/j.linged.2019.03.001

Strong, M., & Baron, W. (2004). An analysis of mentoring conversations with beginning teachers: Suggestions and responses. *Teaching and Teacher Education*, *20*(1), 47–57. https://doi.org/10.1016/j.tate.2003.09.005

ten Have, P. (2007). *Doing conversation analysis.* Sage.

Turhan, B., & Kirkgöz, Y. (2021). A critical and collaborative stance towards retrospective reflection in language teacher education. *European Journal of Teacher Education*, *46*(2), 222–240. https://doi.org/10.1080/02619768.2021.1917545

Vasquez, C. (2004). "Very carefully managed": Advice and suggestions in post-observation meetings. *Linguistics and Education*, *26*, 33–58.

Veen, M., & de la Croix, A. (2016). Collaborative reflection under the microscope: Using conversation analysis to study the transition from case presentation to discussion in GP residents' experience sharing sessions. *Teaching and Learning in Medicine*, *28*(1), 3–14. https://doi.org/10.1080/10401334.2015.1107486

Walsh, S., & Mann, S. (2015). Doing reflective practice: A data-led way forward. *ELT Journal*, *69*(4), 351–362. https://doi.org/10.1093/elt/ccv018

Waring, H. Z. (2013). Two mentor practices that generate teacher reflection without explicit solicitations: Some preliminary considerations. *RELC Journal*, *44*(1), 103–119. https://doi.org/10.1177/0033688212473296

Waring, H. Z. (2014). Mentor invitations for reflection in post-observation conferences: Some preliminary considerations. *Applied Linguistics Review*, *5*(1), 99–123. https://doi.org/10.1515/applirev-2014-0005

Waring, H. Z. (2017). Going general as a resource for doing advising in post-observation conferences in teacher training. *Journal of Pragmatics*, *110*, 20–33. https://doi.org/10.1016/j.pragma.2017.01.009

Wright, T. (2010). Second language teacher education: Review of recent research on practice. *Language Teaching*, *43*(3), 259–296. https://doi.org/10.1017/S0261444810000030

Ulusötesi Öğretmen Adaylarının Epistemik Duruşlarının İşbirlikçi Veri Odaklı Yansıtıcı Diyaloglar İçinde ve Aracılığıyla Değişimi

Öz

Öğretme ve öğrenmeye sosyal odaklı bakış açısını takiben, yansıtıcı düşünme, öğretmen adaylarını yetiştirmede önemli bir pedagojik araç olmuştur. Bu çalışma, Çok kipli Konuşma Çözümlemesi kullanarak bir eğitmen olmaksızın ulusötesi öğretmen adaylarının iş birliğine dayalı veri odaklı yansıtıcı diyaloglarını incelemektedir. Çalışma, iş birliğine dayalı çevrimiçi görev tasarımı, görevlerin gerçek yabancı dil öğrencileri tarafından uygulanması, ve öğrencilerin performansı ve sanal değişim projesi (s=72) deneyimleri üzerine yansıtıcı düşünmeyi içeren bir sanal değişim projesinin ekran kayıtlarından yararlanmaktadır. Çalışmanın odak noktası, öğretmen adaylarının iş birliğine dayalı video aracılı veri odaklı yansıtıcı diyaloglarıdır. Verinin yakından incelenmesi, iş birliğine dayalı yansıtıcı diyaloglar aracılığıyla öğretmen adaylarının (i) uygulamalarına ilişkin farkındalıklarını artırdıklarını ve bilgi ürettiklerini; (ii) teori ve uygulama arasında bağlantı kurduklarını; (iii) epistemik konumlarının daha fazla farkına vardıklarını; (iv) bir sorunu tanımlayıp tarif ettiklerini ve bu sorunlara çözümler bulduklarını göstermektedir (Farrell, 2015). Dolayısıyla, öğrenme ve anlamaya dair temel anlar yansıtıcı düşünme esnasında ve yansıtıcı düşünme yoluyla ortaya konmuştur. Bu çalışma, öğretmen öğrenimi için fırsatlar yaratmak amacıyla ulusötesi öğretmen eğitimi programlarında öğretmen adaylarının yansıtıcı diyaloglarını teşvik etmeye dair çıkarımlarda bulunmaktadır.

Anahtar kelimeler: öğretmen eğitimi, öğretmen öğrenimi, veriye dayalı yansıtıcı düşünme, güç asimetrisi, iş birliği

Appendix

Jefferson (2004) Transcription Convention

[ ]	Overlapping utterances – (beginning [) and (end])
=	Contiguous utterances (or continuation of the same turn)
(0.4)	Represent the tenths of a second between utterances
(.)	Represents a micro-pause (1 tenth of a second or less)
:	Elongation (more colons demonstrate longer stretches of sound)
.	Fall in pitch at the end of an utterance
-	An abrupt stop in articulation
?	Rising in pitch at utterance end (not necessarily a question)
CAPITAL	Loud/forte speech
	Underline letters/words indicate accentuation
↑↓	Marked upstep/downstep in intonation
° °	Surrounds talk that is quieter
hhh	Exhalations
.hhh	Inhalations
he or ha	Laugh particle
(hhh)	Laughter within a word (can also represent audible aspirations)
> <	Surrounds talk that is spoken faster
< >	Surrounds talk that is spoken slower
(( ))	Analyst notes
( )	Approximations of what is heard
$ $	Surrounds 'smile' voice

Mondada (2018) Multimodal Transcription Convention

* *	Gestures and descriptions of embodied actions are delimited between
+ +	two identical symbols (one symbol per participant)
Δ Δ	and are synchronized with corresponding stretches of talk.
*--->	The action described continues across subsequent lines
---->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,,	Action's retraction.
ric	Participant doing the embodied action is identified when (s)he is not the speaker.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a specific symbol showing its position within the turn at talk.

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
2024, Vol. 41-3, No. 1
https://doi.org/10.52597/buje.1411656

Investigating Differential Item Functioning of an International Mathematics Competition Items across Gender Groups

Serkan Arıkan[a]

Abstract

Mathematical problem-solving competitions have existed for over a century. Scholars report the gender gap in these competitions. As a result, it is necessary to determine whether any score difference between gender groups is attributable to a genuine difference or is the result of the exam itself. Thus, the current study specifically examined bias in one of the well-known mathematics competitions: the Kangaroo Mathematics competition. Determining the fairness of Kangaroo mathematics competition items across gender groups is crucial for creating accurate comparisons and avoiding unintended construct irrelevant bias. To examine the bias, Differential Item Functioning (DIF) analyses were conducted using Logistic Regression, Mantel-Haenszel, and Item Response Theory Likelihood Ratio Test DIF detection methods. After a series of investigations, out of 336 items, it was concluded that these mathematics items were free of DIF and bias across the gender groups. Further implications were discussed in detail regarding the validity and bias.

Keywords: DIF, bias, mathematical problem-solving competitions, logistic regression, Mantel-Haenszel, Item Response Theory Likelihood Ratio Test

Article info

Received: 29.12.2023
Revised: 04.03.2024
Accepted: 14.03.2024
Published online: 22.04.2024

Introduction

Mathematical problem-solving competitions date back more than a century. Since then, there have been numerous mathematics competitions that have piqued the interest of students. These mathematics competitions are valuable not only for identifying talented students, but also for encouraging and developing mathematical ability, providing both an opportunity to deal with original problem situations, and a new and high-quality material for classroom teaching and math clubs (de Losada & Taylor, 2022). Compared to school-based exams, these competitions have become attractive as they are external and there is no need to worry about the results (de Losada & Taylor, 2022). Similarly, starting in 1991, Kangaroo mathematics competitions have been attracting more than six million students all around the world each year in more than 70 countries. This annual international mathematical game contest, run by a non-profit organization called the Association Kangourou Sans Frontières, aims to increase the popularization of mathematics in schools, support mathematical education, and increase the joy of mathematics by promoting a positive perception towards mathematics in society (Akveld, Caceres-Duque, Geretschläger, 2020). The competition is run on the same day all around the world. There are six categories according to grades (Pre-Ecolier: Grade 1 and 2; Ecolier: Grade 3 and 4, Benjamin: Grade 5 and 6, Cadet: Grade 7 and 8, Junior: Grade 9 and 10 and Student: Grade 11 and 12). Students in each category answer the same questions, but evaluation and comparison are done within each grade level. In order to develop the test, the candidate items are written and submitted in English by member countries and then selected in the Annual Meeting by teachers, mathematicians, and math educators. Then, selected item sets are translated into target languages by country offices (Akveld, Caceres-Duque, & Geretschläger, 2020). Detailed descriptions of the items used in this competition were provided by scholars (Andritsch et al., 2020; Geretschläger & Donner, 2022).

Each country administers their own test, does the scoring of their students, and announces the results. In the Kangaroo Mathematics competition, cross-country score comparisons are not made. Although the Kangaroo Mathematics competition is not a high-stakes exam, students compete fiercely to be successful. As a result, the comparability of test scores across gender groups and the fairness of items for boys and girls are essential issues, as any comparative assessment should be fair to both groups of pupils (International Test Commission, 2001). Scholars are interested in the gender differences in mathematics performance (Hyde, et al. 2008) and some

[a]*Boğaziçi University, Mathematics and Science Education, serkan.arikan1@bogazici.edu.tr, ORCID: 0000-0001-9610-5496*

To cite: Arıkan, S. (2024). Investigating differential item functioning of an international mathematics competition items across gender groups. *Boğaziçi Unıversity Journal of Education*, *41-3*(1), 53-69. https://doi.org/10.52597/buje.1411656

reported males perform far better than females in competitive environments compared to non-competitive environments (Gneezy et al., 2003; Niederle & Vesterlund, 2010). Applebaum and Leikin (2019) reported that the gender gap is task dependent on mathematics competitions. Thus, it is required to identify whether any score difference between gender groups is due to true differences or stems from the test itself. Consequently, assessing the fairness of Kangaroo Mathematics competition items across gender groups is critical for making reliable comparisons and avoiding unintentional construct-irrelevant bias.

When conducting a comparative study or organizing a competition, it is required to distinguish between the impact and the bias. Impact occurs when one group genuinely has more or less ability on the construct of interest. To assess potential bias in scores, Differential Item Functioning (DIF) detection methods are used to evaluate items. DIF occurs and threatens the comparability of test scores of groups when students who have the same ability level on the construct of interest do not have the same probability of answering the item correctly for a single item (Holland & Thayer, 1988; van de Vijver & Leung, 1997; Zumbo, 2007). Examining and presenting evidence that items are free of DIF is necessary for valid score interpretations. Otherwise, if a test has items showing DIF, observed score differences for specific groups could be due to construct-irrelevant variance rather than the true differences in the ability (He & van de Vijver, 2013). DIF is a statistical term and when an item shows DIF, it means that the item might be biased. As a result, to determine whether an item is biased or not, experts examine the item to determine whether the item functions as biased against one group or not (Allalouf et al., 1999; van de Vijver & Leung, 1997). This expert evaluation is done by investigating the possible sources and causes of bias. DIF can stem from a variety of sources, including inadequate translations, unclear original items, low familiarity/appropriateness of the item content in some cultures/groups, the influence of issues specific to a particular culture/groups, such as nuisance factors or connotations associated with the item wording (van de Vijver & Tanzer, 2004), and contextual factors, such as the socioeconomic status and classroom practices (Zumbo & Gelin, 2005).

There has been a lot of research done on identifying biased items (Berrío et al., 2020). DIF studies with real datasets mainly focus on evaluating items of large-scale assessments such as Programme for International Student Assessment **(**PISA), The Trends in International Mathematics and Science Study (TIMSS), and The National Assessment of Educational Progress (NAEP) (Arikan, 2019; Kankaraš & Moors, 2014; Lyons-Thomas et al., 2014; Reynolds et al., 2022; Roberson & Zumbo, 2019; Zwick & Ercikan, 1989), college admission tests (Dorans, 2013; Stark et al., 2004; Wedman, 2018), job application tests (Stark et al., 2004), and licensing examinations (Rubright et al., 2022). The examination of DIF in large-scale and high-stakes tests is required since test scores are utilized to make critical judgments. Today, fairness in classroom assessment is getting attention (Baniasadi et al., 2023).

Present Study

More people are becoming aware of mathematical competitions and Olympiads, such as the TUBITAK Science Olympiads, The International Mathematical Olympiad (IMO), The American Mathematics Competition (AMC) and the Kangaroo Mathematics Competition. Although there are many worldwide mathematics competitions or mathematics Olympiads, there is a lack of research on the psychometric properties of the tests used in these competitions. There is just one research that focus on investigating gender DIF in the American Mathematics Competition contests administered between 2003 and 2007 (Desjarlais, 2009). The results showed that only two of the 125 items were determined to have non-negligible DIF in accordance with the ETS criteria. The Kangaroo competition is one of the most well-known mathematical competitions, with about 6000000 kids competing annually globally. However, to our knowledge, there is no study investigating DIF in Kangaroo Mathematics competition items. Thus, there is a need to evaluate Kangaroo Mathematics items to understand whether any gender differences in test scores may reflect true differences in the mathematics ability or have bias in test items. Therefore, the current study aims to evaluate Kangaroo Mathematics items in terms of gender DIF. To achieve this goal, Kangaroo Mathematics items that are used in the Turkish version of the competition were tested to investigate whether they contain gender DIF or not. The research question of this study is “Do Kangaroo Mathematics items show DIF across gender groups?”

Method

Participants

Kangaroo Mathematics competition is administered to students from grade 1 to grade 12. The current study participants were students who attended Kangaroo Mathematics competition in Turkey in 2022. DIF analyses were conducted for each grade level separately. The number of students in each grade, gender group mean scores, and effect size of the differences are presented in Table 1. Overall, boys performed better than girls; however, the effect size calculations showed that these differences were small according to Cohen's d (Cohen, 1988), except for grade 11. In grade 11, the difference was medium.

Table 1

Descriptive Statistics of Participants

Grade	Number of Students		Mean Scores		Effect Size
	Boys	Girls	Boys	Girls	
Grade 1	3249	2504	7.43	6.94	0.14*
Grade2	5579	4516	10.35	10.02	0.09*
Grade3	7772	6045	8.90	8.33	0.15*
Grade4	5697	4593	11.52	10.73	0.18*
Grade5	6438	5138	11.97	10.62	0.30*
Grade6	4921	3784	15.14	13.66	0.28*
Grade7	3775	3096	10.74	9.74	0.22*
Grade8	2067	1686	14.01	12.54	0.28*
Grade9	1352	1282	17.34	15.62	0.32*
Grade10	1355	1308	17.71	15.91	0.35*
Grade11	761	622	16.96	14.46	0.43*
Grade12	244	175	16.63	15.49	0.18

* $p < 0.05$

Instrument

Kangaroo Mathematics competition have mathematics items that aim to increase student interest in mathematics and promote mathematical thinking. Every year, the competition uses mathematical items that are proposed by each country. Next, representatives from each country gather to choose and determine the final set of items. Following the selection of the items in English, each nation works with its scientific committee to translate and adapt the items. Items that are inappropriate for the national curriculum may be substituted by this committee.

Each grade level consists of 24 or 30 multiple choice questions (from grade 1 to grade 4, 24 items; from grade 5 to grade 12, 30 items) and there are three item categories according to the anticipated difficulty of the items (3-point, 4-point, and 5-point-problems). Grade 1 and 2, grade 3 and 4, grade 5 and 6, grade 7 and 8, grade 9 and 10, grade 11 and grade 12 took the same test but no comparison is made across grade levels. The content dimensions of the test are Numbers, Geometry, Combinatorics, and Algebra. One example item for 7th and 8th graders is provided in Figure 1 (Akveld, Caceres-Duque, Nieto Said & Sánchez Lamoneda, 2020).

Figure 1

Kangaroo Mathematics Example Item

Cadet, 2019-14. Alan, Bella, Claire, Dora, and Erik met at a party and shook hands exactly once with everyone they already knew. Alan shook hands once, Bella shook hands twice, Claire shook hands three times and Dora shook hands four times. How many times did Erik shake hands?

Ⓐ 1; Ⓑ 2; Ⓒ 3; Ⓓ 4; Ⓔ 0.

Data Analysis

There are many methods for DIF detection (Berrío et al., 2020). As with other statistical tests, DIF results can have Type 1 error (false positive). Using multiple DIF detection methods allows researchers a triangulation of the results. Thus, in the current study, three different DIF detection methods were used. To obtain more reliable results, an item flagged as showing DIF in at least two methods was regarded to have DIF across gender groups. The DIF detection methods selected in the current study were logistic regression (LR), Mantel-Haenszel (MH), and the Item Response Theory Likelihood Ratio Test (IRT-LR). LR and MH are based on Classical Test Theory where the ability match is done by total scores and these methods apply parametric and non-parametric models, respectively. IRT-LR is based on Item Response Theory and the ability match is done by latent traits. LR could identify nonuniform DIF in addition to uniform DIF. In Uniform DIF, one group constantly gains an advantage throughout all ability level. Nevertheless, with nonuniform DIF, the conditional dependency shifts and reverses at various points on the ability level rather than providing the reference group with a constant advantage over the ability continuum. Thus, these three DIF detection methods are considered to represent a wide range of statistical procedures. From grades 1 through 12, a total of 336 mathematics items were assessed for both boys and girls in terms of DIF.

In the logistic regression DIF detection method, the total score, grouping variable, and interaction term are included in the model hierarchically. Changes in R^2 values as specified below are interpreted as evidence for uniform DIF or nonuniform DIF (Zumbo, 1999). There are two mainly accepted criteria for DIF detection: Zumbo and Thomas (1997) proposed that ΔR^2 higher than 0.130 indicates moderate DIF and higher than 0.260 indicates large DIF; Jodoin and Gierl (2001) proposed lower values to detect DIF such as ΔR^2 higher than 0.035 indicates moderate DIF and higher than 0.070 indicates large DIF. In the current study, in Model 1, only the total score; in Model 2, the total score and gender; and in Model 3, the total score, gender, and their interaction were used as predictors, and ΔR^2 value of 0.035 was chosen as the threshold for detecting more items. SPSS 27 was used to conduct logistic regression DIF analysis.

The Mantel-Haenszel DIF detection method creates K two-by-two contingency tables, where K represents the number of discrete total score intervals to match group abilities. For each score interval, the expected and observed ratios are calculated and the difference is tested by the chi-square method (Holland & Thayer, 1986). As the chi-square method is highly affected by the large sample size, the MH D-DIF index is proposed for the evaluation where a negative value indicates that an item favors the reference group over the focal group (Holland & Thayer, 1988). Educational Testing Service (ETS) proposed a criterion to flag DIF items: The MH D-DIF index between 1.00 and 1.50 indicates moderate DIF and higher than 1.50 indicates large DIF (Zieky, 1993). The current study used an MH D-DIF index value of 1.00 as the criterion for detecting more items. DIFAS 5.0 (Penfield, 2005) was used for MH DIF detection analysis.

Item Response Theory Likelihood Ratio Test evaluates whether an item has the same IRT item parameters for the reference and focal groups. In the IRT-LR detection procedure, first, all parameters are forced to be equal for both groups (the compact model). Then, parameters are allowed to vary for the studied item (the augmented model), and the difference between these models are compared with a likelihood ratio (LR) test $G^2 = - 2LL_C - (-2LL_A)$ where -2LL denotes the negative two log-likelihood of the compact and augmented models). This difference follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters estimates between the models (Thissen et al., 1993). As the significance of the chi-square test is easily affected by sample size and is difficult to interpret (Stark et al., 2004), the effect size of the b parameter difference could be used to flag DIF items. Steinberg and Thissen (2006) stated that a b parameter difference of 0.50 or larger represents a large effect. As the sample size was very large in the current study, the b parameter differences were estimated, and items were flagged accordingly. IRTLRDIF v2.0b software (Thissen, 2001) was used for the computations.

If an item is detected to have DIF, then expert opinions are crucial to decide whether the item has bias or not. The important technique for determining test item characteristics (such as content, format, context, or language) that may lead to DIF is expert (or judgmental) review, which involves reviewing items by people who are aware about student learning and may have linguistic or cultural experience (Roth et. al, 2013). Thus, in the current study, items flagged as having DIF were evaluated by three experts. Three experts were included in the evaluation committee: a math teacher, a measurement expert, and an academician with expertise in measurement and evaluation. They all had experience on developing and evaluating mathematics items.

Results

Reliability and Unidimensionality of the Instrument

The reliability of the test scores was evaluated using Cronbach's alpha internal consistency coefficient (See Table 2). For each grade level, internal consistency values were ranging from 0.69 to 0.89. These values suggested that the internal consistency of items in each grade level was acceptable. As the study focused on gender differences, these values were computed independently for each gender group. Cronbach's alpha values of gender groups were quite close. For grade levels 3, 4 and 7, Cronbach's Alpha values were relatively higher for boys. This discrepancy suggested that certain items may have had a poorer correlation with girls' overall scores. Whether or not these items create bias was analyzed in the next section.

Evaluation of unidimensionality according to the ratio-of-first-to-second eigenvalues-greater-than-three rule (Slocum-Gori & Zumbo, 2011) and minimum average partial (MAP) showed that for most of the grade levels there is one general factor. MAP is one of the suggested way of deciding number of factor to retain (Zwick & Velicer, 1986) and estimated by the following application (https://afarukkilic.shinyapps.io/Factor_Analysis_For_All_FAFA).

Table 2

Reliability of Test Scores and Unidimensionality

Grade	Cronbach's alpha (All)	Cronbach's alpha (Boys)	Cronbach's alpha (Girls)	First and second eigenvalues	the ratio-of-first-to-second eigenvalues	Minimum Average Partial (MAP)
1	0.69	0.70	0.68	3.107-1.326	2.34	1
2	0.70	0.71	0.69	3.170-1.179	2.69	1
3	0.71	0.73	0.66	3.341-1.288	2.59	1
4	0.77	0.79	0.74	4.042-1.333	3.03	1
5	0.77	0.78	0.74	4.085-1.387	2.95	1
6	0.82	0.82	0.81	4.960-1.330	3.73	1
7	0.75	0.77	0.72	3.933-1.438	2.74	1
8	0.81	0.82	0.80	4.912-1.550	3.17	1
9	0.85	0.85	0.84	5.901-1.723	3.42	2
10	0.85	0.85	0.83	5.753-1.764	3.26	2
11	0.87	0.88	0.84	6.331-1.542	4.11	1
12	0.89	0.89	0.88	7.236-1.432	5.05	1

DIF Results

DIF analyses were conducted using Logistic Regression, Mantel-Haenszel, and Item Response Theory Likelihood Ratio Test DIF detection methods (See Table 3 through Table 8). An item that was flagged by at least two methods was considered as showing DIF. The LR DIF results were reported based on ΔR^2. In terms of both uniform and non-uniform DIF, all ΔR^2 values were lower than 0.035. Thus, it was concluded that the LR DIF method did not detect any item showing DIF in any grade level. The MH results were reported based on the MH D-DIF index. The results revealed that none of the items had an MH D-DIF value of more than 1.00. Thus, according to the MH DIF detection method, none of the items was flagged in any grade level. In the IRT-LR detection procedure, the b parameter differences were reported and the difference of 0.50 or larger was used as DIF detection. For Grade 1 and Grade 9, none of the items; for Grade 2, items 1, 6, 11; for Grade 3, items 1, 2, 3, 11; for Grade 4, items 2, 4, 18; for Grade 5, items 1, 3, 4, 27; for Grade 6, items 1, 4, 5, 27, 29; for Grade 7, item 15; for Grade 8, item 4; for Grade 10, items 1 and 15; for Grade 11, item 11, and for Grade 12, items 3, 4,10 were flagged as having DIF.

Table 3

Grade 1 and Grade 2 DIF Results

	Grade1				Grade2			
Item No	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*
1	0.002	0.000	-0.300	-0.070	0.002	0.000	-0.302	**0.720***
2	0.000	0.003	-0.066	-0.460	0.000	0.000	-0.112	-0.270
3	0.001	0.000	0.096	0.150	0.000	0.001	-0.054	-0.360
4	0.005	0.000	0.297	0.340	0.003	0.000	0.231	0.340
5	0.000	0.000	-0.003	0.040	0.001	0.000	-0.148	-0.050
6	0.000	0.000	0.043	-0.480	0.000	0.003	0.042	**-0.610***
7	0.005	0.001	-0.362	-0.240	0.010	0.000	-0.482	-0.270
8	0.001	0.000	-0.073	0.230	0.001	0.000	-0.111	-0.120
9	0.000	0.001	-0.105	-0.300	0.002	0.000	-0.185	-0.100
10	0.000	0.000	-0.068	-0.180	0.001	0.000	0.154	0.380
11	0.000	0.001	0.023	0.070	0.001	0.001	0.168	**0.500***
12	0.002	0.001	-0.194	-0.300	0.003	0.000	-0.240	-0.200
13	0.002	0.000	-0.181	-0.210	0.003	0.001	-0.254	-0.170
14	0.001	0.000	-0.094	-0.160	0.001	0.000	-0.162	-0.260
15	0.003	0.001	-0.255	0.110	0.001	0.000	-0.113	-0.080
16	0.000	0.000	0.019	0.420	0.000	0.000	-0.041	-0.230
17	0.001	0.000	0.068	-0.060	0.001	0.000	0.104	0.020
18	0.002	0.001	0.197	-0.010	0.001	0.000	0.118	-0.020
19	0.000	0.000	0.085	-0.100	0.001	0.000	-0.166	-0.030
20	0.000	0.001	0.035	0.200	0.000	0.000	0.080	0.000
21	0.000	0.000	0.119	0.000	0.005	0.000	0.323	0.090
22	0.002	0.000	0.177	0.110	0.002	0.000	0.212	0.080
23	0.001	0.000	0.143	-0.040	0.007	0.000	0.373	0.010
24	0.003	0.000	0.247	0.130	0.001	0.000	0.134	0.030

Note: * indicates the item shows DIF

Table 4

Grade 3 and Grade 4 DIF Results

	Grade3				Grade4			
Item No	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*
1	0.003	0.000	0.239	**0.550***	0.004	0.000	0.298	0.420
2	0.026	0.001	0.773	**0.510***	0.033	0.001	0.838	**0.730***
3	0.007	0.001	-0.361	**-0.510***	0.006	0.000	-0.326	-0.200
4	0.003	0.001	-0.192	-0.430	0.003	0.001	-0.222	**-2.140***
5	0.001	0.000	0.066	0.010	0.000	0.000	0.036	0.120
6	0.001	0.000	-0.137	-0.060	0.001	0.000	-0.091	-0.020
7	0.002	0.000	0.268	0.150	0.002	0.000	0.226	0.200
8	0.001	0.000	-0.084	-0.020	0.002	0.000	-0.173	-0.250
9	0.000	0.000	-0.055	0.000	0.000	0.000	-0.002	0.120
10	0.000	0.000	0.056	0.090	0.000	0.001	0.054	0.030
11	0.012	0.001	-0.451	**-0.640***	0.010	0.000	-0.425	-0.360
12	0.001	0.000	-0.128	-0.160	0.002	0.000	-0.217	-0.260
13	0.000	0.000	0.065	-0.420	0.001	0.001	0.124	-0.030
14	0.000	0.001	0.087	0.160	0.000	0.000	-0.065	0.010
15	0.001	0.000	-0.094	-0.160	0.002	0.000	-0.232	-0.160
16	0.000	0.002	0.052	-0.110	0.001	0.000	0.068	0.020
17	0.000	0.000	-0.051	-0.150	0.001	0.000	-0.177	-0.170
18	0.005	0.000	0.319	0.470	0.014	0.000	0.507	**0.730***
19	0.000	0.000	-0.067	-0.150	0.000	0.000	-0.084	-0.060
20	0.001	0.001	0.054	0.060	0.000	0.001	0.029	-0.100
21	0.001	0.000	0.160	-0.020	0.002	0.000	0.187	0.060
22	0.000	0.000	-0.011	-0.170	0.001	0.000	-0.111	-0.010
23	0.001	0.000	-0.081	-0.270	0.000	0.001	-0.012	-0.230
24	0.000	0.000	-0.023	-0.210	0.001	0.000	0.093	0.180

Note: * indicates the item shows DIF

Table 5

Grade 5 and Grade 6 DIF Results

	Grade5				Grade6			
Item No	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*
1	0.008	0.000	-0.470	**-0.620***	0.012	0.000	-0.651	**-0.750***
2	0.003	0.000	0.223	0.170	0.001	0.000	0.099	0.080
3	0.011	0.001	0.464	**0.570***	0.012	0.000	0.483	0.480
4	0.007	0.000	-0.372	**-0.550***	0.005	0.000	-0.349	**-0.550***
5	0.017	0.000	0.650	0.350	0.015	0.000	0.672	**0.580***
6	0.003	0.000	0.277	0.090	0.002	0.000	0.262	0.150
7	0.003	0.006	-0.185	-0.320	0.005	0.004	-0.259	-0.230
8	0.001	0.001	0.058	0.440	0.005	0.002	0.335	0.330
9	0.005	0.000	-0.350	-0.310	0.002	0.001	-0.212	-0.190
10	0.001	0.001	-0.149	-0.030	0.000	0.000	0.054	-0.020
11	0.000	0.001	-0.070	-0.190	0.001	0.000	-0.139	-0.060
12	0.000	0.000	-0.065	-0.230	0.001	0.000	-0.075	-0.360
13	0.003	0.000	-0.240	-0.290	0.002	0.000	-0.201	-0.160
14	0.006	0.000	0.361	0.280	0.011	0.000	0.513	0.190
15	0.000	0.000	-0.059	-0.110	0.002	0.001	-0.193	-0.110
16	0.000	0.000	-0.006	-0.090	0.000	0.000	-0.040	-0.040
17	0.002	0.000	0.158	0.210	0.001	0.000	0.137	0.150
18	0.013	0.000	-0.499	-0.490	0.015	0.000	-0.517	-0.350
19	0.001	0.000	0.123	0.070	0.001	0.001	0.154	0.040
20	0.001	0.000	-0.104	-0.100	0.000	0.000	-0.051	-0.040
21	0.000	0.000	0.076	0.020	0.000	0.000	0.059	-0.160
22	0.001	0.000	-0.093	-0.360	0.001	0.001	-0.191	-0.360
23	0.001	0.000	-0.115	-0.150	0.001	0.000	-0.147	-0.190
24	0.003	0.001	0.263	0.200	0.002	0.000	0.196	0.130
25	0.001	0.000	-0.120	-0.140	0.002	0.000	-0.213	-0.250
26	0.001	0.000	-0.092	0.040	0.000	0.001	-0.039	0.030
27	0.001	0.000	-0.113	**-1.100***	0.004	0.000	-0.241	**-1.200***
28	0.000	0.001	0.146	0.270	0.000	0.001	0.124	0.270
29	0.001	0.001	0.201	0.000	0.002	0.001	0.225	**1.050***
30	0.000	0.000	0.014	0.070	0.002	0.000	0.202	-0.050

Note: * indicates the item shows DIF

Table 6

Grade 7 and Grade 8 DIF Results

	Grade7				Grade8			
Item No	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*
1	0.000	0.001	0.087	0.170	0.002	0.000	0.180	.250
2	0.002	0.001	-0.226	-0.210	0.003	0.004	-0.270	-.180
3	0.005	0.000	0.330	0.310	0.003	0.000	0.288	.240
4	0.000	0.000	-0.042	-0.300	0.000	0.004	0.082	**1.180***
5	0.003	0.002	-0.282	-0.300	0.002	0.003	-0.247	-.180
6	0.000	0.000	-0.014	0.160	0.000	0.001	-0.011	0.340
7	0.001	0.000	-0.152	-0.230	0.000	0.001	-0.061	-0.270
8	0.006	0.001	-0.327	-0.310	0.000	0.000	-0.076	-0.040
9	0.005	0.001	-0.332	0.000	0.005	0.001	-0.306	-0.120
10	0.000	0.000	0.056	0.000	0.000	0.001	-0.058	-0.070
11	0.001	0.000	0.134	0.170	0.001	0.000	0.150	0.150
12	0.000	0.000	0.023	-0.040	0.000	0.000	-0.039	-0.030
13	0.001	0.000	0.156	0.090	0.001	0.000	0.115	0.040
14	0.001	0.000	0.128	-0.080	0.000	0.000	-0.040	-0.220
15	0.022	0.003	0.699	**0.780***	0.031	0.000	0.874	0.470
16	0.000	0.000	0.065	0.030	0.000	0.000	0.013	0.100
17	0.005	0.000	-0.343	-0.200	0.004	0.000	-0.292	-0.230
18	0.000	0.001	-0.044	-0.100	0.001	0.001	0.200	-0.150
19	0.000	0.000	0.109	-0.190	0.001	0.002	0.170	-0.060
20	0.005	0.000	0.346	0.050	0.001	0.000	0.137	0.040

	Grade7				Grade8			
Item No	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*
21	0.000	0.001	0.057	-0.110	0.000	0.000	-0.010	0.020
22	0.001	0.000	-0.145	-0.210	0.003	0.000	-0.221	-0.160
23	0.000	0.000	0.034	0.020	0.001	0.000	-0.076	0.100
24	0.005	0.003	0.360	0.120	0.004	0.006	0.337	0.310
25	0.001	0.000	0.094	-0.320	0.000	0.000	-0.016	-0.190
26	0.001	0.001	-0.158	-0.130	0.000	0.000	-0.058	0.050
27	0.002	0.001	-0.207	-0.240	0.002	0.001	-0.218	-0.100
28	0.000	0.001	-0.031	0.080	0.000	0.000	-0.010	0.040
29	0.001	0.000	-0.071	-0.400	0.002	0.000	-0.200	-0.280
30	0.001	0.001	0.199	0.400	0.004	0.000	0.292	0.060

Note: * indicates the item shows DIF

Table 7

Grade 9 and Grade 10 DIF Results

	Grade9				Grade10			
Item No	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*
1	0.007	0.002	-0.438	0.250	0.010	0.001	-0.506	**1.280***
2	0.000	0.001	-0.130	-0.110	0.001	0.000	-0.185	-0.060
3	0.000	0.000	-0.060	0.080	0.003	0.000	-0.245	-0.020
4	0.000	0.001	-0.004	-0.050	0.000	0.001	-0.036	0.130
5	0.000	0.000	-0.031	-0.010	0.000	0.001	-0.025	-0.090
6	0.000	0.001	0.112	0.010	0.000	0.000	-0.081	-0.050
7	0.001	0.000	-0.159	-0.020	0.002	0.000	-0.192	-0.260
8	0.005	0.000	-0.399	-0.270	0.006	0.000	-0.402	-0.210
9	0.002	0.000	0.253	0.110	0.002	0.000	0.227	0.300
10	0.000	0.001	0.167	0.030	0.000	0.001	0.083	-0.060
11	0.000	0.000	-0.044	-0.050	0.000	0.000	-0.059	0.010
12	0.001	0.001	0.150	-0.200	0.002	0.000	0.194	0.190
13	0.000	0.000	-0.109	-0.080	0.000	0.000	-0.059	-0.090
14	0.000	0.000	0.052	0.030	0.000	0.000	-0.041	0.050
15	0.000	0.002	-0.065	-0.460	0.001	0.004	-0.137	**0.670***
16	0.001	0.000	-0.162	-0.160	0.001	0.001	-0.173	-0.150
17	0.016	0.001	0.708	0.440	0.007	0.000	0.433	0.290
18	0.000	0.001	0.128	0.100	0.001	0.000	-0.145	-0.160
19	0.000	0.003	-0.113	0.270	0.000	0.000	0.076	-0.030
20	NA	NA	NA	NA	NA	NA	NA	NA
21	0.001	0.000	0.118	0.080	0.000	0.000	0.088	-0.070
22	0.002	0.005	0.234	-0.120	0.012	0.001	0.633	0.020
23	0.001	0.000	0.089	0.060	0.005	0.000	0.372	0.190
24	0.001	0.001	0.175	-0.190	0.001	0.000	0.147	0.160
25	0.000	0.000	-0.122	-0.070	0.000	0.000	-0.043	-0.150
26	0.000	0.000	0.012	0.010	0.000	0.000	-0.034	-0.140
27	0.000	0.000	-0.043	-0.040	0.000	0.000	-0.040	-0.090
28	0.004	0.001	-0.367	-0.210	0.003	0.001	-0.284	-0.270
29	0.007	0.003	-0.435	-0.410	0.002	0.001	-0.211	-0.170
30	0.005	0.003	0.385	-0.400	0.010	0.000	0.520	0.140

Note: * indicates the item shows DIF; NA: The item was canceled due to printing issues.

Table 8

Grade 11 and Grade 12 DIF Results

	Grade11				Grade12			
Item No	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*	*LR* ΔR^2 M_2-M_1	*LR* ΔR^2 M_3-M_2	*ΔMH*	*Δb*
1	0.000	0.001	0.047	0.240	0.000	0.000	-0.025	-0.060
2	0.000	0.000	-0.133	-0.020	0.007	0.000	-0.527	-0.160
3	0.017	0.006	-0.598	0.140	0.004	0.002	0.873	**1.070***
4	0.006	0.001	-0.355	-0.210	0.003	0.004	0.291	**0.750***
5	NA	NA	NA	NA	NA	NA	NA	NA

	Grade11				Grade12			
Item No	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*	*LR ΔR²* M_2-M_1	*LR ΔR²* M_3-M_2	*ΔMH*	*Δb*
6	0.003	0.000	-0.199	-0.240	0.001	0.002	0.137	0.150
7	0.000	0.002	-0.001	0.200	0.001	0.002	0.000	0.240
8	0.001	0.001	0.178	0.320	0.004	0.000	0.277	0.400
9	0.005	0.001	-0.388	-0.190	0.005	0.001	-0.568	-0.260
10	0.004	0.003	-0.294	0.030	0.006	0.008	-0.346	**-0.540***
11	0.026	0.000	0.774	**0.970***	0.006	0.000	0.237	0.470
12	0.004	0.000	-0.342	-0.240	0.000	0.015	0.030	-0.250
13	0.002	0.002	-0.246	-0.180	0.010	0.005	-0.684	-0.080
14	0.001	0.000	0.125	0.270	0.010	0.000	0.449	0.390
15	0.009	0.001	0.413	0.410	0.023	0.000	0.708	0.330
16	0.018	0.001	0.659	0.210	0.014	0.001	0.545	0.140
17	0.000	0.001	0.086	-0.030	0.004	0.004	0.409	0.380
18	0.010	0.000	-0.557	-0.330	0.009	0.006	-0.602	-0.150
19	0.001	0.000	0.097	-0.050	0.001	0.001	-0.283	-0.030
20	0.003	0.001	-0.275	-0.080	0.008	0.000	-0.608	-0.200
21	0.001	0.000	-0.141	-0.070	0.000	0.002	-0.042	0.050
22	0.000	0.001	0.107	-0.140	0.014	0.001	0.430	0.080
23	0.000	0.000	0.127	0.230	0.003	0.000	-0.323	-0.020
24	0.000	0.000	-0.166	-0.060	0.003	0.001	0.283	0.270
25	0.001	0.000	0.094	-0.180	0.006	0.002	-0.513	-0.040
26	0.001	0.001	0.184	0.190	0.001	0.001	-0.129	-0.120
27	0.000	0.001	-0.023	0.090	0.001	0.000	0.192	0.090
28	0.000	0.001	-0.072	-0.010	0.001	0.009	-0.135	0.020
29	NA	NA	NA	NA	NA	NA	NA	NA
30	0.000	0.000	-0.016	-0.020	0.000	0.003	-0.119	-0.300

Note: * indicates the item shows DIF; NA: The item was canceled due to printing issues.

Evaluation of DIF Results

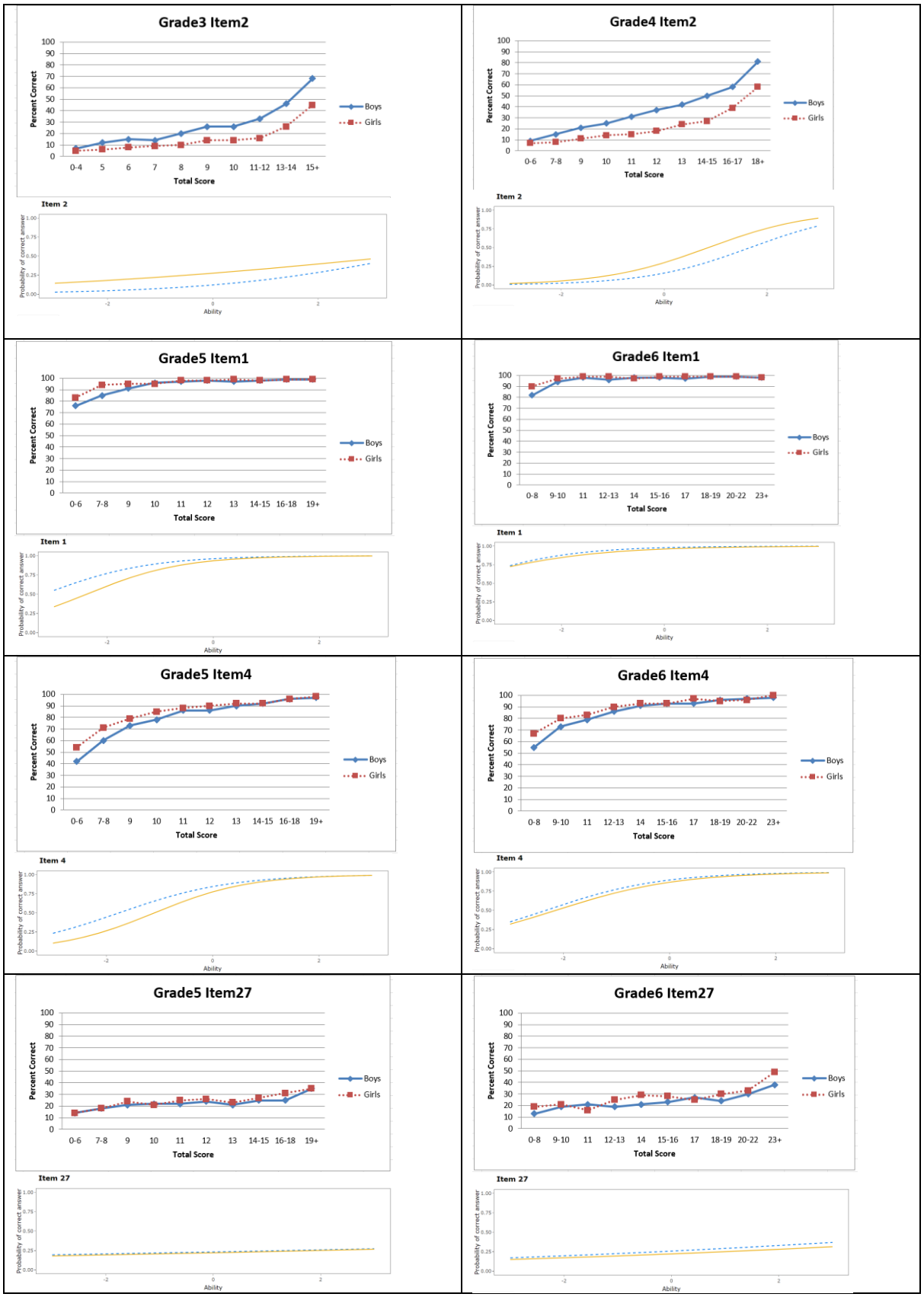
Overall, out of 336 items, LR DIF and MH flagged none of the items whereas IRT-LR flagged 27 items as having DIF with 16 items favoring boys, and 11 items favoring girls. As none of the items was flagged by two or more DIF detection methods, it was concluded that 2022 Kangaroo Mathematics items were free of uniform or nonuniform DIF and bias across the gender groups.

However, an item flagged as showing DIF in both grade levels (as each item was administered to two grade levels) was worth examining. Thus, items 2 in Grades 3 and 4; and items 1, 4 and 27 in Grades 5 and 6 were evaluated by the experts to determine if these items contain any bias (these four items are presented in the appendix). Additionally, to provide more information for this examination, percentages of correct responses and ICC's for gender groups were drawn (see Figure 2). In percentage correct graphs, the x-axis represents the total score groups created according to deciles, and the y-axis represents the percentage correct responses of gender groups. In ICCs x-axis represents the ability score estimated by IRT, and the y-axis represents the probability of correct answer of gender groups. These graphs indicated that, controlling student ability, the curves did not differ seriously.

Following the process of detecting DIF items empirically, there is a post hoc procedure to examine DIF items judgmentally to detect the possible causes of DIF, if any. First, the direction of DIF was determined. Item 2 in Grades 3 and 4 favored boys; whereas, items 1, 4 and 27 in Grades 5 and 6 favored girls. Then, these four items were evaluated by three experts. The experts were given the items and asked if they anticipated any DIF. The DIF detection results such as the direction of DIF items were also presented. Possible sources of DIF were highlighted to make experts to focus on those sources. The experts claimed that the item wording or pictures of the items were free of gender bias. Thus, all experts concluded that these four items did not contain any gender-specific content to create gender bias.

Figure 2

Percentages of correct responses and ICC's for gender groups



Evaluation of Gender Mean Score Differences Excluding DIF Items

As it was shown in Table 1, boys perform better than girls; however, these differences were small according to Cohen's d (Cohen, 1988), except for grade 11. Effect size and the standardized mean difference allow for comparing the difference between groups without being affected by sample size (Field, 2013). In this part, to evaluate the negative consequences associated with the presence of DIF, the original effect sizes and the effect sizes excluding DIF items were compared (see Table 9). Overall, the effect sizes of the gender differences were very similar with or without DIF items. Thus, it was concluded that these DIF items did not produce any biased consequences on scores.

Table 9. Effect Sizes of with and without DIF items

Grade	Effect Size (original)	Effect Size (excluding DIF)
Grade 1	0.14*	0.14*
Grade2	0.09*	0.08*
Grade3	0.15*	0.15*
Grade4	0.18*	0.13*
Grade5	0.30*	0.29*
Grade6	0.28*	0.28*
Grade7	0.22*	0.20*
Grade8	0.28*	0.28*
Grade9	0.32*	0.32*
Grade10	0.35*	0.36*
Grade11	0.43*	0.41*
Grade12	0.18	0.18

* $p < 0.05$

Discussion

Mathematical problem-solving competitions provide opportunities for mathematicians and mathematics educators to collaborate, create tests together and conduct joint research (de Losada & Taylor, 2022). It is considered that scholars from the field of educational measurement will involve in these tests in the near future; and by looking through the lenses of measurement specialists, the research on the psychometric properties of these competition tests will increase. Focusing on a well-known mathematical problem-solving competition, the research papers on Kangaroo Mathematics competitions are mainly related to the evaluation of content (Jiang, & Xiong, 2021), description of the competition, and problems (Akveld, Caceres-Duque, & Geretschläger, 2020; Akveld, Caceres-Duque, Nieto Said & Sánchez Lamoneda, 2020), test-wiseness strategies (Donner et al., 2021), gender gap and task relationship (Applebaum & Leikin, 2019), the effect of teacher gender (Escardibul & Mora, 2013), and comparing student performances in competitions and classroom tests (Mellroth, 2015). There is a lack of research on the psychometric properties of mathematical problem-solving competitions. As a result, the current study addressed the issue of gender equity through the lenses of psychometry and examined Kangaroo mathematics competition items for DIF and bias.

The preliminary results showed that boys had higher mathematics scores in the competition than girls, and these differences were mainly small according to Cohen's d (Cohen, 1988). When there is a difference between two groups, the difference could be due to either an impact or bias. To understand the nature of the differences, DIF analyses were conducted. Out of 336 items, none of the items displayed DIF on two DIF detection methods (LR DIF and MH) and 27 items displayed DIF on the IRT-LR method, 16 items favoring boys and 11 items favoring girls. Thus, it was concluded that there was not an issue of DIF across gender groups because no items were detected by two or more DIF detection methods. To have a deeper investigation, items that displayed DIF in one method in both grade levels were identified. These four items were examined based on the direction of DIF (three out of four items favored girls), percentages of correct responses for gender groups, and expert evaluations. As a result of this examination, no evidence for bias was found. Finally, the effect sizes with or without DIF items (omitting 27 items flagged by IRT-LR) were compared and no change in effect sizes was observed. Overall, based all of these detailed investigations, it was concluded that there was no gender bias on Kangaroo Mathematics competition items. Thus, the small sized difference observed between gender groups could not be explained by bias.

Examining DIF is not only related to individual test scores, but also to the educational quality of the assessment instruments by providing validity evidence (Berrío et al., 2020). The standards for educational and psychological testing require test publishers to report the evidence of reliability and validity (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Evaluating the possible bias and providing evidence for bias-free assessment scores are essential for valid score interpretations. Increasing fairness in testing increases validity in scores and helps minimize the construct-irrelevant variances (ETS, 2022). Thus, like other large-scale assessments, mathematical competitions are also needed to provide fair and valid assessment across subgroups. By evaluating bias in one of the major mathematical competitions, the current study is expected to lead to other studies that focus on the psychometric properties of mathematics competitions.

It is vital to highlight that fairness begins with the preparation of the test plan and the development of the items, because introducing unfair content or construct-irrelevant variance for some groups of examinees may result in bias (ETS, 2022). For instance, considering gender bias, the inclusion of items based on soccer-related calculations on the league table could be a source of bias. Similarly, having items related to metro systems of metropolitan cities may create bias across students located in rural and urban regions. As validity is related to the interpretation and consequences of test scores (Messick, 1989), an item that could have a detrimental consequence on a certain group needs to be avoided. Thus, test developers and test reviewers need to be cautious about the issue of bias. Additionally, test takers need to be knowledgeable about this issue and demand bias-free assessments.

The current study concentrated on a gender bias in data from a single country. It is suggested that future studies compare international mathematical competitions cross-culturally. Evaluating the reliability and validity of these competitions across cultures and evaluating measurement invariance at the test-level and DIF at the item-level would add valuable information to the literature. International large-scale assessments such as PISA and TIMSS have provided an assessment framework before the administration and technical reports after the administration. Similarly, reporting psychometric findings of international mathematics competitions on an annual basis as technical reports would also increase the interest of more scholars and stakeholders throughout the world. In addition to gender groups and country-level comparisons, there are other groups to consider for the fairness of test results. ETS (2022) recommended that special attention should be paid to the following groups on any assessment: age, appearance, citizenship status, disability, ethnicity, gender, national or regional origin, native language, race, religion, sexual orientation and socioeconomic status. Thus, future studies could examine DIF across these groups in mathematics competitions.

This mathematics competition was also given to first graders. First graders are still developing their reading skills. Thus, even though they volunteered to participate in this competition, their score might have construct-irrelevant variance. Given that the first graders' scores had an adequate Cronbach's alpha, it appears that there was no issue regarding their age. Nonetheless, studying first graders' experiences in these competitions might yield important data for subsequent studies.

References

Akveld, M., Caceres-Duque, L. F., & Geretschläger, R. (2020). Math Kangaroo. *Mathematics Competitions*, *33*(2), 48–66.

Akveld, M., Caceres-Duque, L. F., Nieto Said, J. H., & Sánchez Lamoneda, R. (2020). The Math Kangaroo Competition. *Espacio Matemático*, *1*(2), 74–91.

Allalouf, A., Hambleton, R. K., & Sireci, S. G. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement*, *36*(3), 185–198. https://doi.org/10.1111/j.1745-3984.1999.tb00553.x

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, D.C: American Educational Research Association.

Andritsch, L., Hauke, E., & Kelz, J. (2020). How to create and solve: Analysis of items from the Mathematical Kangaroo from two perspectives. In R. Geretschläger (Ed.), *Engaging young students in mathematics through competitions—World perspectives and practices, Vol. II: Mathematics competitions and how they relate to research, teaching and motivation* (pp. 117–136). World Scientific.

Applebaum, M., & Leikin, R. (2019). Girls' performance in the Kangaroo contest. In M. Nolte (Ed.), *Including the Highly Gifted and Creative Students–Current Ideas and Future Directions-Proceedings of the 11th International Conference on Mathematical Creativity and Giftedness (mcg 11)*, (pp. 87–94). Hamburg, Germany.

Arikan, S. (2019). Are Differentially Functioning Mathematics Items Reason of Low Achievement of Turkish Students in PISA 2015? Journal of Measurement and Evaluation in Education and Psychology, 10(1), 49–67. https://doi:10.21031/epod.466860

Baniasadi, A., Salehi, K., Khodaie, E., Bagheri Noaparast, K., & Izanloo, B. (2023). Fairness in classroom assessment: A systematic review. *The Asia-Pacific Education Researcher*, *32*, 91–109. https://doi.org/10.1007/s40299-021-00636-z

Berrío, Á. I., Gomez-Benito, J., & Arias-Patiño, E. M. (2020). Developments and trends in research on methods of detecting differential item functioning. *Educational Research Review*, *31*, 100340. https://doi.org/10.1016/j.edurev.2020.100340

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, (2nd ed.). Hillsdale, NJ: Erlbaum.

de Losada, M. F., & Taylor, P. J. (2022). Perspectives on mathematics competitions and their relationship with mathematics education. *ZDM–Mathematics Education*, *54*(5), 941–959. https://doi.org/10.1007/s11858-022-01404-z

Desjarlais, M. A. (2009). *Gender differences on the American Mathematics Competition AMC 8 contest*. The University of Nebraska-Lincoln.

Donner, L., Kelz, J., Stipsits, E., & Stuhlpfarrer, D. (2021). Which test-wiseness based strategies are used by Austrian winners of the Mathematical Kangaroo?. *Mathematics Competitions*, *34*(1), 88–101.

Dorans, N. J. (2013). ETS contributions to the quantitative assessment of item, test, and score fairness. *ETS Research Report Series*, *2013*(2), i–38.

Escardibul, J. O., & Mora, T. (2013). Teacher gender and student performance in Mathematics. Evidence from Catalonia (Spain). *Journal of Education and Training Studies*, *1*(1), 39–46.

ETS. (2022). ETS guidelines for developing fair tests and communications. https://www.ets.org/content/dam/ets-org/pdfs/about/fair-tests-and-communications.pdf

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Geretschläger, R., & Donner, L. (2022). Writing and choosing problems for a popular high school mathematics competition. *ZDM–Mathematics Education*, *54*(5), 971–982. https://doi.org/10.1007/s11858-022-01351-9

Gneezy, U., Muriel, N., & Aldo, R. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, *118*(3), 1049–1074. https://doi.org/10.1162/00335530360698496

He, J., & van de Vijver, F. J. R. (2013). Methodological issues in cross-cultural studies in educational psychology. In G. A. D. Liem & A. B. I. Bernardo (Eds.), *Advancing cross-cultural perspectives on educational psychology: A festschrift for Dennis McInerney (*pp. 39–56). Charlotte, NC: Information Age Publishing.

Holland, P. W., & Thayer, D. T. (1986). *Differential item functioning and the Mantel-Haenszel procedure* (ETS Research Report No. RR-86-31). Princeton, NJ: ETS.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity*, (pp.129–145). Hillsdale, N.J.: Erlbaum.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. https://doi.org/10.1126/science.1160364

International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, *1*(2), 93–114. https://doi.org/10.1207/S15327574IJT0102_1

Jiang, P., & Xiong, B. (2021, April). Analyze the quality of Math Kangaroo problems with a content analysis. In *Journal of Physics: Conference Series* (Vol. 1875, No. 1, p. 012015). IOP Publishing.

Jodoin, M. G., & Gierl, M. J. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, *14*(4), 329–349. https://doi.org/10.1207/S15324818AME1404_2

Kankaraš, M., & Moors, G. (2014). Analysis of cross-cultural comparability of PISA 2009 scores. *Journal of Cross-Cultural Psychology*, *45*(3), 381–399. https://doi.org/10.1177/0022022113511297

Lyons-Thomas, J., Sandilands, D. D., & Ercikan, K. (2014). Gender differential item functioning in Mathematics in four international jurisdictions. *Education &Science*, *39*(172), 20–32.

Mellroth, E. (2015). Problem solving competency and the mathematical kangaroo. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME9, 4-8 February 2015)* (pp. 1095–1096). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME. *https://hal.science/CERME9/public/CERME9_Proceedings_2015.pdf*

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd edition, pp.13–103). New York, NY: Macmillan.

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in Math test scores: The Role of Competition. *Journal of Economic Perspectives*, *24*(2), 129–144. https://doi.org/10.1257/jep.24.2.129

Penfield, R. D. (2005). DIFAS: Differential Item Functioning Analysis System. *Applied Psychological Measurement*, *29*, 150–151. https://doi.org/10.1177/0146621603260686

Reynolds, K., Khorramdel, L., & von Davier, M. (2022). Can students' attitudes towards mathematics and science be compared across countries? Evidence from measurement invariance modeling in TIMSS 2019. *Studies in Educational Evaluation*, *74*, 101169. https://doi.org/10.1016/j.stueduc.2022.101169

Roberson, N. D., & Zumbo, B. D. (2019). Migration background in PISA's measure of social belonging: Using a diffractive lens to interpret multi-method DIF studies. *International Journal of Testing*, *19*(4), 363–389. https://doi.org/10.1080/15305058.2019.1632316

Roth, W.-M., Oliveri, M. E., Sandilands, D., Lyons-Thomas, J., & Ercikan, K. (2013). Investigating sources of differential item functioning using expert think-aloud protocols. *International Journal of Science Education*, *35*, 546–576. https://doi.org/10.1080/09500693.2012.721572

Rubright, J. D., Jodoin, M., Woodward, S., & Barone, M. A. (2022). Differential item functioning analysis of United States medical licensing examination step 1 items. *Academic Medicine*, *97*(5), 718–722. https://doi.org/10.1097/ACM.0000000000004567

Slocum-Gori, S. L., & Zumbo, B. D. (2011). Assessing the unidimensionality of psychological scales: Using multiple criteria from factor analysis. *Social Indicators Research*, *102*, 443–461. https://doi.org/10.1007/s11205-010-9682-8

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2004). Examining the effects of differential item (functioning and differential) test functioning on selection decisions: When are statistically significant effects practically important?. *Journal of Applied Psychology*, *89*(3), 497–508. https://doi.org/10.1037/0021-9010.89.3.497

Steinberg, L., & Thissen, D. (2006). Using effect sizes for research reporting: examples using item response theory to analyze differential item functioning. *Psychological methods*, *11*(4), 402–415. https://doi.org/10.1037/1082-989X.11.4.402

Thissen, D. (2001). IRTLRDIF v2.0b: *Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning* [Documentation for computer program]. L.L. Thurstone Psychometric Laboratory, University of North Carolina at Chapel Hill.

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P. W. Holland & H. Wainer (Eds.), *Differential Item Functioning, (*pp.67–113). Mahwah, NJ: Erlbaum.

van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis of comparative research*. Thousand Oaks, CA: Sage.

van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *54*(2), 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Wedman, J. (2018). Reasons for gender-related differential item functioning in a college admissions test. *Scandinavian Journal of Educational Research*, *62*(6), 959–970. https://doi.org/10.1080/00313831.2017.1402365

Zieky, M. (1993). Practical questions in the use of DIF statistics in item development. In P. W. Holland & H. Wainer (Eds.), *Differential Item Functioning, (*337–364). Hillsdale, NJ: Lawrence Erlbaum.

Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores.* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, *4*(2), 223–233. https://doi.org/10.1080/15434300701375832

Zumbo, B. D., & Gelin, M. N. (2005). A matter of test bias in educational policy research: Bringing the context into picture by investigating sociological/community moderated (or mediated) test and item bias. *Journal of Educational Research & Policy Studies*, *5*(1), 1–23.

Zumbo, B. D., & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF*. Prince George, Canada: Edgeworth Laboratory for Quantitative Behavioral Science, University of Northern British Columbia.

Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP History assessment. *Journal of Educational Measurement, 26*, 55–66. https://doi.org/10.1111/j.1745-3984.1989.tb00318.x

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432

Uluslararası Matematik Yarışması Maddelerinin Cinsiyet Gruplarına Göre Değişen Madde Fonksiyonu (DMF) açısından İncelenmesi

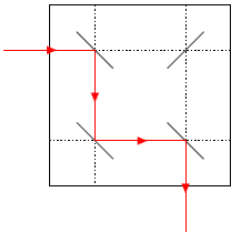
Öz

Matematiksel problem çözme yarışmaları bir asırdan fazla bir süredir uygulanmaktadır. Araştırmalar bu yarışmalarda kızlar ve erkekler arasında başarı farkı olduğunu göstermektedir. Cinsiyet grupları arasındaki puan farkının gerçekte var olan bir farklılıktan mı yoksa sınavdaki sorulardan mı kaynaklandığının belirlenmesi değerli bilgiler sunacaktır. Bu nedenle bu çalışma pek çok öğrencinin katıldığı matematik yarışmalarından biri olan Uluslararası Kanguru Matematik yarışmasındaki madde yanlılığını incelemektedir. Maddelerin Değişen Madde Fonksiyonu (DMF) içerme durumlarını incelemek için Lojistik Regresyon, Mantel-Haenszel ve Madde Tepki Kuramı Olabilirlik Oranı Testi kullanılmıştır. Analizlerden elde edilen bulgulara göre toplam 336 maddeden oluşan bu matematik testlerinin DMF içermediği, dolayısıyla maddelerin cinsiyet grupları arasında yanlılık yaratmadığı sonucuna varılmıştır. Makalede geçerlilik ve yanlılıkla ilgili diğer çıkarımlar ayrıntılı olarak tartışılmıştır.

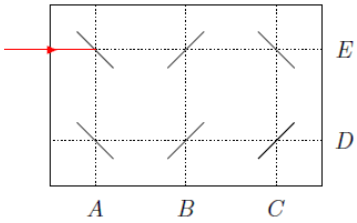
Anahtar kelimeler: DMF, yanlılık, uluslararası matematik yarışmaları, lojistik regresyon, Mantel-Haenszel ve Madde Tepki Kuramı Olabilirlik Oranı Testi

Appendix A

Grade 3-4, Item 2



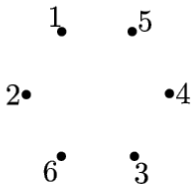
2. The mirrors reflect the laser beam like this:
Where does this laser beam end?



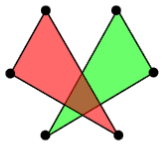
(**A**) A (**B**) B (**C**) C (**D**) D (**E**) E

Grade 5-6, Item 1

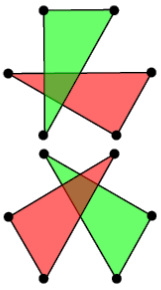
There are six points numbered as shown. We create two triangles, one by connecting the points with even numbers, the other one by connecting the points with odd numbers. Which of the five figures do we get?



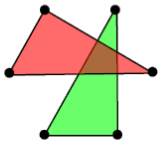
(**A**)



(**B**)

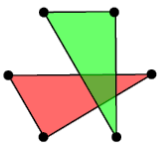


(**C**)



(**D**)

(**E**)



Grade 5-6, Item 4

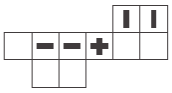
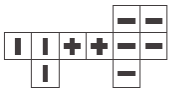
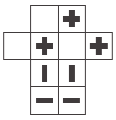
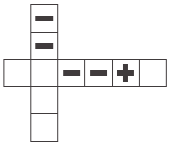
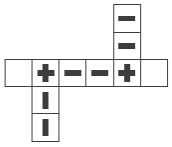
Cistercian numerals were used in the early thirteenth century. Any integer from 1 to 99 can be represented by a single glyph, combining the glyphs below.

So 24 looks like  , 81 looks like  and 93 looks like  . How does 45 look like?

(A)  (B)  (C)  (D)  (E) 

Grade 5-6, Item 27

Which of the following nets cannot be folded into the solid  ?

(A)  (B)  (C)  (D)  (E) 

BOĞAZİÇİ UNIVERSITY JOURNAL OF EDUCATION
2024, Vol. 41-3, No. 1
https://doi.org/10.52597/buje.1441617

Navigating Disciplinary Dialogues: An Exploratory Analysis of Dialogic Stance

Nur Yigitoglu Aptoula[a] and Betul Bal Gezegin[b]

Abstract

A number of studies have focused on the analyses of theses, articles, and book reviews as academic genres, yet there are also other genres that have recently appeared in academic journals and need to be analyzed. One of these genres is the so-called "disciplinary dialogues," which are relatively short academic texts written by experts to discuss a specific issue rather than reporting research. This study focuses on authors' interaction in disciplinary dialogues. The data for the study is a specialized corpus comprised of 37 disciplinary dialogues published in a journal in the Applied Linguistics field. Discourse analysis was employed. The findings illuminate a rich array of writing strategies employed by authors to establish a dialogic stance, such as referring to other scholars/studies, asking rhetorical questions, explicitly showing disagreement/agreement, using hedging devices, using evaluative and reflective language, and suggesting future research ideas. The paper ends with the pedagogical and research implications of the findings.

Keywords: academic journals, author interaction, disciplinary dialogues, dialogic stance, corpus

Article info

Received: 22.02.2024
Revised: 12.03.2024
Accepted: 15.03.2024
Published online: 22.04.2024

Introduction

Researchers have been extensively addressing the theme of interaction in academic discourse over the past few decades. The question of how writers interact with their readers and the language they use to do so have been discussed under various concepts, such as metadiscourse (Crismore, 1989; Hyland & Tse, 2004), attitude (Halliday, 1994), epistemic modality (Hyland, 1998), appraisal (Martin, 2000; White, 2003), stance (Biber & Finegan, 1989) to name a few (Hyland, 2005). To reveal how interaction is constructed and performed through the use of language in different academic disciplines across a range of academic genres has been found valuable in the field of English for Specific Purposes, particularly in English for Academic Purposes.

Disciplinary dialogues are short academic texts published in academic journals and written to clarify a new concept, present perspectives on that concept, and discuss its future potential and implications. As its name suggests, "disciplinary dialogues" are anticipated to create a discourse by discussing a major topic in the field. A disciplinary dialogue constitutes a text that is written by scholars who engage in debate through written texts, given the impracticality of face-to-face discussions. Unlike research articles, these dialogues do not aim to report the results of research. As opinion papers, however, these pieces reflect the author's thoughts on a certain issue by covering this specific issue with a critical perspective. They provide the readers with an in-depth analysis of the issue by putting it in a larger context. Disciplinary dialogues are usually written on an issue or paper that was previously published, and, in some instances, several disciplinary dialogues are written on the same paper. For this reason, we believe that it is important to investigate disciplinary dialogues as a newly-emerged yet still unexplored genre within the field of second language writing.

While disciplinary dialogues seem to be a genre specific to Applied Linguistics, in other disciplines, similar types of writing under other labels, such as "colloquia" or "symposia" can also be found. Very few, if any, studies have focused on these genres. It is also possible to find similar "responses" as genres that are common,

[a]*Corresponding author, Boğaziçi University, Department of Foreign Language Education, nuryigitoglu@gmail.com, ORCID: 0000-0002-9039-6672*
[b]*Samsun Ondokuz Mayıs University, Department of English Language Teaching, betulbal@gmail.com, ORCID: 0000-0001-9818-9347*

To cite: Yigitoglu Aptoula, N. & Bal Gezegin, B. (2024). Navigating disciplinary dialogues: An exploratory analysis of dialogic stance. *Boğaziçi University Journal of Education*, *41-3*(1), 71-79. https://doi.org/10.52597/buje.1441617

and in these pieces, authors respond to a published piece. Similarly, there are "position papers," in which authors purposefully take a (perhaps controversial) stand on an issue, often to provoke discussion. The disciplinary dialogues section, on the other hand, includes both of these later types of texts, featuring a lengthier position paper and shorter responses from various authors. In fact, the disciplinary dialogues section includes two different genres - the position paper being one and the responses being another. That is why it is novel and still remains an underexplored genre within the field of Applied Linguistics.

Given the complexity of this specific genre, one can argue that disciplinary dialogues are research-professional genres that experienced scholars are invited to write. Atkinson (2012), when commenting on the initiation of the section of this specific genre in an Applied Linguistics journal, specifically describes the disciplinary dialogues section as focused on "debate," with a position paper detailing an "argument" and "responses" from others. He also notes that the disciplinary dialogues section focuses on "a major topic under discussion in the field." (p.1) As Atkinson noted, he felt “the debate needed to continue,” and this idea has remained as the ultimate goal of the scholars writing disciplinary dialogues. Several themes have been debated in disciplinary sections throughout the years: L2 writing expertise, multimodal composing, argumentation, replication in second language writing, contrastive rhetoric, and plagiarism, to name a few. These topics remain subject to debate in the field of Applied Linguistics, but one should also note that the disciplinary dialogues contribute helpful perspectives on the field that will then generate a larger discussion in the field about these issues and establish it as an important component of applied linguistics scholarship.

In order to push the area forward and to gain a deeper understanding of the norms of academic writing within Applied Linguistics, the norms of disciplinary dialogues as a genre need to be addressed. Although it could be argued that this is a genre of experienced scholars, we believe that this genre should be dealt with in university writing courses for a number of reasons. Firstly, disciplinary dialogues as a genre require synthesis and dialogic communication, which are skills needed for critical thinking skills. In addition, graduate students are not always expected to produce argumentative essays or research papers in all their content courses; they might be expected to respond to content or reading materials in different ways. Disciplinary dialogues as a research-professional genre require a deeper level synthesis and contributes to the arguments in the field. Graduate students typically transition from being primarily knowledge consumers to becoming knowledge producers as they advance in their academic and professional careers. However, even as they generate new knowledge, they continue to consume and integrate existing knowledge to inform their work. For this reason, we believe it is important to help our students “unpack textual norms” (Hyland, 2018, p. 383) and transcend their ingrained presumptions about the serious tone and formulaic ways of academic writing. Teaching disciplinary dialogues, in particular, may promote graduate students’ critical thinking development as they engage in this genre and critically analyze and evaluate the existing research. Also, as new comers to the academic discourse community, graduate students may be more aware of disciplinary norms and arguments within the field. Lastly, although contributing to intellectual discourse by participating in disciplinary dialogues may be an expert writing task, it is important for graduate students to become aware of outgoing debates within the field and engage in debates surrounding hot topics related to their expertise.

Creating a Dialogic Stance in Disciplinary Dialogues

The incorporation of interaction into texts has been an issue of all academic genres, yet some genres need to be more dialogic due to their generic features. Although it is common to come across the use of the word “dialogic” or expressions like “dialogic pedagogy” and “dialogic teaching” mostly in spoken contexts, it needs to be discussed in written discourse as well. The authors of certain texts are expected to use more critical, expressive, evaluative, and interactive language due to the academic text type they are writing.

Dialogic communication in this paper is used as a way of interaction that the authors create while they “detail their views in print,” “provide written responses on focus articles,” be in “meta-level discussions” (Atkinson, 2012, p.283). With these responsibilities in mind, the authors of disciplinary dialogues position themselves in a triangulated communication of authors themselves, readers, and focus article authors. We believe how this interaction among the three parties is realized needs to be investigated to understand this academic genre better.

With this goal in mind, this paper aims to investigate how a dialogic stance is achieved in disciplinary dialogues. In this paper, we follow Hyland’s (2008) interactional model of voice that consists of stance and engagement and his perspective that authors reflect their presence in their texts and construct a relationship with their readers. He states that “every successful academic text displays the writer’s awareness of both its readers

and its consequences" (p. 6). Authors of academic texts present their awareness of their authors and manage their interaction with them by relying on certain strategies. Rather than a quantitative approach to identifying patterns of authorial voice, this study adopts a macro level and holistic qualitative analysis of authorial voice in "disciplinary dialogues," which have not been investigated as a genre before.

Method

Selection of Texts

In order to examine dialogic stance in disciplinary dialogues, 37 target texts published in an Applied Linguistics journal were collected. The corpus under consideration encompasses a total of 44,934 words. Notably, each individual text within this corpus exhibits a similar word count and adheres to the established criteria, obviating the necessity for the exclusion of any specific target texts from its composition. It should be noted that since this is a relatively new genre, the number of journals publishing disciplinary dialogues is limited. The first disciplinary dialogue was published in 2012 in the *Journal of Second Language Writing*, and since then, the journal has kept on publishing and promoting this new academic genre. To ensure that similar texts are not missed to be included in this special corpus, a list of prestigious journals in the field was checked to determine whether they include texts that have been written and published with similar purposes. It is found that it is not a common academic genre published in most journals. In addition to a manual check of the journals and personal recommendations, queries in the search engines with different keywords such as "disciplinary dialogues," "short communication," "short report," and "scholarly discussions" were conducted to ensure that similar texts that were written for the same purpose with a similar structure are not missed. The inclusion criteria of texts are as follows:

- Question and answer types of texts are excluded.
- Texts that report a research study are excluded.
- A similar text length is ensured (too short and too long texts are excluded)
- Only single-author dialogue texts are included.
- Meta-analyses are excluded.

Data Analysis

Discourse analysis was undertaken to analyze and interpret the gathered data. A discourse can be defined as a system of texts that brings objects into being (Parker, 1992). The researchers carried out a manual and cyclical analysis of each text to identify the strategies used by the authors to create the dialogic stance. In this discourse analysis, the researchers individually analyzed the possible patterns that indicate the dialogic and interactive tone of the authors. Each strategy found was tagged and checked by two researchers for reliability. The data analysis was carried out by following these steps:

1. Searching for depictions and cues related to dialogic or interactive markers within the texts.
2. Annotating and categorizing the interpretation of these depictions, with a specific emphasis on discerning the functions of the descriptions and their associated linguistic features within the text.
3. Bringing together all relevant instances from each text and formulating a comprehensive compilation of the recurring strategies identified.

Findings

The analysis of 37 disciplinary dialogues revealed that authors use common strategies to reach a dialogic tone and interact with the readers as well as the authors of focus articles. These can be listed as follows:

1. Citing other sources/scholars
2. Asking rhetorical questions
3. Showing agreement/disagreement
4. Using hedging devices

5. Suggesting future research ideas
6. Using reflective language
7. Using evaluative language

Citing Other Sources/Scholars

It has been observed that authors of disciplinary dialogues frequently refer to other scholars to support their evaluations. By citing the works of other researchers, they strengthen their dialogic tone and evaluations. Some examples can be found below:

"As highlighted by Smith et al. (2018), the role of environmental factors cannot be overlooked in this context."

"An early, and important work is Derewianka (1991) exploring how texts work in which she unpacks features of systemic functional language description..."

As can be seen from the examples, the authors cite other scholars to make their dialogic tone stronger in the disciplinary dialogues.

Asking Rhetorical Questions

As the name suggests, disciplinary dialogues provide the readers with the author's disciplinary expertise in the form of a dialogue. A common strategy used by the authors to build this dialogue is integrating questions in their texts. This strategy helps to engage with the readers in the form of dialogue and prompts reflection.

"For instance, in what specific aspects of source use do students and experts differ most and least?"

"Then, how can (should) we establish multimodal literacy research..."

"But why is that important?"

The examples above show these rhetorical questions in disciplinary dialogues serve as a way to build their own arguments and to engage the readers in their disciplinary dialogues.

Showing Agreement/Disagreement

The most frequent feature that helps to create a dialogic stance in the texts is showing agreement or disagreement. Particularly, agreements were found in all of the texts, and it is observed that expressing agreement with the authors of the articles discussed signals a willingness to find common ground and fosters a cooperative tone in the dialogue. Unlike other academic genres, explicit expressions of agreement and disagreement indicate that there are generic features of disciplinary dialogues, and the authors of these texts foster a conversational tone and create a scholarly exchange, which is the objective of writing these texts. Some examples of agreement can be found below:

"I agree with Pecorari that avoiding plagiarism is largely about learning."

"Let me first confirm the authors' point that, with some caveats and cautions, quasi-replications are needed..."

"I agree with Li that the greatest collective strength and impact of contrastive/ cross-cultural/intercultural studies have been in the comparative analyses of texts..."

"Therefore, I agree wholeheartedly with her assertion that we should apply ourselves to disseminating our research and to educating our institutional colleagues."

As can be seen from the examples above, the authors use agreements to create dialogic engagement. The authors also use disagreement to build their dialogic tone better. Below are some examples of disagreements found in the disciplinary dialogues:

"Porte and Richards seem to take a narrow view of L2 writing research..."

> "Oddly, however, in attempting to mix Confucian and Aristotelian approaches in her writing, Professor Li is supporting the idea that culture is complex and hybrid..."

> "Space limitations only allow me to rebut her accusation that I fail to provide evidence in making my arguments."

It should be noted that disagreements and agreements appear together most of the time, as seen in the following examples:

> "I am not as convinced as she is that L2 instructors are behind the curve on this, but I agree this is an important challenge for all language and literacy educators."

> "Despite our many disagreements, I concur with Professor Li…

Using Hedging Devices

The authors tend to use hedging devices to soften their tone. Such devices are varied, and in disciplinary dialogues, they were identified as modal verbs, certain verbs, nouns, adverbs, and adjectives. Here are some examples:

> "I believe that when we explore..."

> "However, it is pertinent to stress that…"

> "The findings suggest a possible correlation between X and Y, though further research is warranted to confirm these tentative conclusions."

> "Contrary to this view, we tend to think that the traditional incarnation of CR has been forced to go through a conceptual paradigm shift…"

As can be seen from the examples above, the authors use a variety of hedging devices to form their arguments in the disciplinary dialogues. These devices seem to help the authors create a dialogic stance and stimulate further discussions on the topic introduced and/or responded in the disciplinary dialogue.

Suggesting Future Research Ideas

The analysis of dialogic disciplinary dialogues has revealed that one of the features that contribute to creating a dialogic stance is the way the authors interact with the readers by drawing attention to possible future research ideas. By doing so, the authors open up space for continued conversation and motivate the researchers to think about potential directions for future studies. Some examples of such uses can be found below:

> "Future research could delve into the impact that expert teachers have on nonexpert teachers, shedding light on the role of mentoring in the ongoing development of teacher expertise."

> "Future research needs to pay adequate attention to classroom-based assessment issues by looking into how teachers and students conceptualize argumentative tasks, and also into scoring procedures used for judging resulting performance."

> "One of the promising areas that future researchers can focus on is the impact that expert teachers have on nonexpert teachers."

These examples illustrate that the authors of disciplinary dialogues usually conclude with implications for future research in an attempt to shed light on the areas why future research needs to know more about the topic of the disciplinary dialogue and what can help move the discipline forward. This discussion helps the authors to build upon the continuation of the topic of the disciplinary dialogue in future research studies.

Using Reflective Language

The authors of disciplinary dialogues were observed to use reflective language, and common phrases, including the personal pronoun "I," were detected in almost all of the texts. The reflective language can create a more inclusive and participatory tone.

> "I am keenly aware that argumentation also plays an important civic role."
>
> "Then, as I read further into the paper, I could see that my own pondering on these issues were, in essence, echoing Hirvela and Belcher's proposal for the concept of adaptive expertise."
>
> "I have been deliberately provocative in this essay because…"
>
> "My position on this issue has been clear and consistent since around 1995 when I started my effort to help establish the disciplinary identity of second language writing."

The examples above show that the first-person pronoun *I* is frequently used in this genre. This makes sense because disciplinary dialogues typically consist of a focus paper that presents the author's (or authors') thoughts on the topic, followed by a series of responses by other authors. Given that the genre showcases the author's expertise on the topic of the disciplinary dialogue, the genre requires the presence of the author within the text and, thus, the use of the pronoun *I*.

Using Evaluative Language

Interpersonality and interactivity are observed to be essential features of dialogic academic texts, which are highly evaluative in nature. Evaluation can be defined as ''expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about'' (Thompson & Hunston, 2000, p. 5). Evaluation in the analyzed texts appeared to be positive, negative, and both positive and negative evaluations. The evaluative acts were identified based on the definition proposed by Suarez-Tejerina (2006):

> "We have counted as an evaluative act any structural unit, irrespective of its lexico-grammatical configuration, that contains both the (sub)aspect commented upon and what is said about it, or what action is recommended" (p. 153)

The authors use positive evaluative language in the disciplinary dialogues to highlight their appreciation of the cited works. Some examples of this kind of positive evaluative language can be found below:

> "I have admired Xiaoming Li's research and writing ever since I read her dissertation,"
>
> "DePalma and Ringer argue for a very diverse, multidimensional, and multidisciplinary framework."
>
> "With their contributions, these authors provide useful pointers to guide further discussion and inquiry."

In some other dialogues, the authors use negative evaluative language to highlight their ideas in contrast to the cited works or a related issue. Here are some examples of such negative evaluative language found in disciplinary dialogues:

> "While I can understand the desire to call it a discipline, declaring it a discipline does not make it so."
>
> "But from my perspective, it is quite irrelevant why they have plagiarized."
>
> Porte and Richards seem to take a narrow view of L2 writing research, their few examples overly focusing on studies of feedback and teacher-student alignment.

While disciplinary dialogues include examples of negative and positive evaluations, findings also revealed that disciplinary dialogues include a combination of positive - negative evaluative language instances. In these examples, the positive evaluation helps the authors underline the niche found in the discipline. Below are some examples of negative-positive evaluations.

> I appreciate her effort to reconstruct my position on the disciplinarity of second language writing, but it made me realize that I have not yet had the chance to fully articulate some of the fundamental concepts underlying my work on disciplinarity.
>
> While some corrective feedback studies have indicated that revision can assist language learning (Chandler, 2003; Ferris, 2006; Shintani et al., 2014), clear evidence of the usefulness of revision itself is still very limited.

Discussion

This short report presents the language used in disciplinary dialogues by expert writers to form the interaction between the readers, the authors of the focused articles, and themselves as the authors of disciplinary dialogues. We believe that the exploration of this academic discourse will be a valuable contribution to the field of ESP, particularly EAP. As evidenced by the insights gleaned from this study, the deliberate choices the authors make and the writing strategies they employ in this specialized genre play a significant role in establishing a dialogic stance. Unveiling what the expert writers do to achieve their goal of discussing the target articles yielded the use of evaluative language. Employing evaluative language seemed to enrich academic communication and provide a platform for reader engagement and creating a dialogue. Exploring the shades of evaluative language within disciplinary dialogues can guide both novice and expert writers in honing their writing skills, enhancing the effectiveness of their communication within the academic discourse. Evaluation is an important yet neglected part of teaching academic writing. Most academic writing teaching books focus on the genre requirements as well as the language required in various parts of those genres. We think, however, that getting graduate students to create their own dialogues with examples like the ones provided here in this short report will make them think more carefully about what they write about others' views. These findings are in line with previous literature which reported that each genre exhibited specific characteristics (Biber & Finegan, 1989).

Pedagogically, the implications of these findings are of significance. Consistent with previous research on metadiscourse and academic writing, this study underlines the importance of understanding the conventions of disciplinary dialogues, informing both research and pedagogy in the field of applied linguistics and beyond. White (2003), for example, also discussed that metadiscourse plays a crucial role in helping writers to engage with their audience and negotiate their stance effectively and highlighted the importance of metadiscourse awareness in academic writing instruction, particularly for novice writers who may struggle with the conventions of disciplinary discourse.

Genre-specific instruction, therefore, is needed, particularly for graduate students who need the tools to write more effectively and to enhance their writing expertise in their academic disciplines. Therefore, integrating the identified writing strategies into writing courses could help those students. Teachers of academic writing for graduate students, for example, could use these examples as a starting point or model sentences in their classes. In these activities, they could ask their students to write their own disciplinary dialogue papers on a topic or paper of their choice. Similarly, for those academic writing classes that include students from different disciplines, teachers could make students analyze disciplinary dialogues from different disciplines to make their students more aware of disciplinary-specific uses of evaluative language used in this genre. After the students write their disciplinary dialogues, they can peer review each other's disciplinary dialogues following a guideline. An example set of questions for such a guideline is provided below:

- Does the disciplinary dialogue present a critical challenge to one or more aspects of the focal article, arguing for a position other than that taken in the focal article?
- Does it include a dialogic stance for each/one/some of the following?:
    - Citing other sources/scholars
    - Asking rhetorical questions
    - Showing agreement/disagreement
    - Using hedging devices
    - Suggesting future research ideas
    - Using reflective language
    - Using evaluative language
- Does it maintain a respectful tone critical to ideas or conclusions but not authors?
- Does it make a clear take-home message?

- Does it include a short title that emphasizes your key message?
- Does it use only essential citations?

In addition, looking at how scholarly authors debate in writing in disciplinary dialogues can be valuable in raising graduate students' awareness of how authors make different rhetorical choices across rhetorical contexts. For example, graduate students could compare how writers disagree in disciplinary dialogues and consider why they make these different choices and what the effects are. Such analysis can help highlight for students the broad range of writing choices that writers choose from and how those choices are connected to rhetorical aims.

Conclusion

The data of this short report is a specialized corpus from an Applied Linguistics journal. The findings are, therefore, limited to the texts obtained from this journal. However, further investigations are needed to show whether the observed writing strategies are specific to this discipline or if they resonate across different disciplines. Comparative studies exploring variations in disciplinary dialogues can be conducted to reach a comprehensive understanding of academic interaction. In addition, as the disciplinary dialogues genre remains unexplored, future research studies could focus on the moves of the genre of disciplinary dialogues. We also believe that there is an interaction between writers' expertise and writers' use of evaluation in disciplinary dialogue sections. Future research studies could investigate this interaction in disciplinary dialogues. Lastly, we propose these analyses should be integrated with writers' views on their evaluative language use. For this reason, future researchers could also interview writers of disciplinary dialogues to tap into their cognitions regarding their own use of evaluative language in disciplinary dialogues.

References

Atkinson, D. (2012). Disciplinary dialogues: Introduction. *Journal of Second Language Writing*, 21, 283.

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text- Interdisciplinary Journal for the Study of Discourse*, *9*(1), 93-124.

Crismore, A. (1989). Talking with readers: Metadiscourse as rhetorical act. *New directions in discourse processing, 14*, 147-175.

Halliday, M. A. K. (1994). *An introduction to functional grammar.* Routledge.

Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.

Hyland, K. (2005). *Metadiscourse*. Continuum.

Hyland K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*. *1*(1), 5-22.

Hyland, K. (2018). Sympathy for the devil? A Defense of EAP. *Language Teaching, 51*(3), 383-399. https://doi.org/10.1017/S0261444818000101

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied linguistics*, *25*(2), 156-177.

Martin, J. R. (2000). Beyond exchange: Appraisal systems in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142-175). Oxford University Press.

Parker, I. (1992). *Discourse dynamics*. Routledge.

Suarez-Tejerina, L. (2006). *Modes of evaluation and rhetorical patterns: a contrastive study of English and Spanish book reviews* [Unpublished doctoral dissertation]. Universidad de Leo´n, Spain.

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In G. Thompson, & S. Hunston (Eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse* (pp. 1-27). Oxford University Press.

White, P. R. R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text & Talk*, *23*(2), 259-284.

Disipliner Diyaloglar: Diyaloğa Özgü Tutumun Keşifsel Bir Analizi

Öz

Bir dizi çalışma akademik türler olarak tez, makale ve kitap incelemelerinin analizine odaklanmıştır, ancak akademik dergilerde son zamanlarda ortaya çıkan ve analiz edilmesi gereken başka türler de vardır. Bu türlerden biri de "disipliner diyaloglar" olarak adlandırılan ve araştırma raporlamaktan ziyade belirli bir konuyu tartışmak üzere uzmanlar tarafından yazılan görece kısa akademik metinlerdir. Bu çalışma, yazarların disipliner diyaloglardaki etkileşimlerine odaklanmaktadır. Çalışmanın verisini, Uygulamalı Dilbilim alanındaki bir dergide yayınlanan 37 disipliner diyalogdan oluşan özel bir derlem oluşturmaktadır. Analiz yöntemi olarak söylem analizi kullanılmıştır. Bulgular, yazarlar tarafından diyalojik bir duruş oluşturmak için kullanılan, diğer akademisyenlere/çalışmalara atıfta bulunma, retorik sorular sorma, uzlaşma/anlaşmazlığı açıkça gösterme, kaçınma (yumuşatma) ifadeleri kullanma, değerlendirici ve yansıtıcı dil kullanma ve gelecekteki araştırma fikirlerini önerme gibi zengin bir dizi yazma stratejisini aydınlatmaktadır. Makale, bulguların pedagojik ve araştırma çıkarımlarıyla sona ermektedir.

Anahtar kelimeler: akademik dergiler, yazar etkileşimi, disipliner diyaloglar, diyalojik duruş, derlem

Boğaziçi University
Journal of Education

Boğaziçi University Journal of Education (BUJE) publishes scientific articles that contribute to the field of education in Turkey and in the world. The Journal publishes refereed articles on theoretical or applied research in the following fields: Psychological services in education, guidance and psychological counseling; early childhood or pre-school education; special education; adult education; educational administration, supervision, philosophy, history and policies; science and math education; foreign language education and applied linguistics; measurement, testing and research techniques in education; educational technology; program development and evaluation in education.

Basic Rules to be Followed in Preparing and Submitting Manuscripts to Boğaziçi University Journal of Education

Manuscripts submitted to the Journal should not have been published or submitted for publication review elsewhere. They could be a research report, a systematic literature review or an original opinion article discussing a new idea or model.

The standards prepared by the American Psychological Association (APA) have been adopted in journal publications. Please refer to the APA 7 publication manual for the rules to be followed in the articles submitted to the journal. Articles should be uploaded to the Dergi Park system at https://dergipark.org.tr/tr/pub/buje.

The rules to be followed in submitting the article are listed below.

The languages of the Journal are Turkish and English. The title of the article should not exceed 20 words, the entire text, excluding figures, tables, and references, should be between 4500-7500 words, and all pages, including the first page, should be numbered in the upper right corner. In addition to the abstract given in the language of the main text, another abstract should be given in the other language. These abstracts should be written on separate pages, with 100-150 words for applied research and 75-120 words for theoretical articles and literature reviews. Below the article abstract, 3 to 5 keywords, which can indicate the main subject of the article and are widely accepted in the scientific language of the field, should be provided.

Throughout the text, a font like Times or Arial (10 points) that allows both Turkish and English characters to be easily typed should be selected, and the same font should be used throughout the text.

To maintain the integrity and fairness of the peer review process, it is essential that submitted manuscripts undergo a thorough anonymization process. Authors are kindly requested to adhere to the following guidelines when preparing their manuscripts for submission:

1. Remove Author Identifying Information: Ensure that all author-specific information is removed from the manuscript file, including names, affiliations, contact details, and any other identifying information in the main body of the text and in headers or footers.

2. Anonymize References: When citing your own work within the manuscript, replace author names and details with "Author" and the publication year to maintain anonymity. For example, replace "Smith et al. (2020)" with "Author (2020)."

3. Blinded Title Page: Create a separate title page that includes the manuscript title but excludes author names and affiliations. This title page should be uploaded as a separate document during the submission process.

4. Anonymous File Names: Ensure that the file names of all submitted documents (including the manuscript file and any supplementary materials) do not contain any author names or identifying information.

5. Remove Acknowledgments: Any acknowledgments that could potentially reveal author identities should be omitted from the manuscript during the anonymization process.

6. Check Citations in Tables and Figures: Review tables and figures to make sure that any author-related information is removed or replaced with anonymized labels.

7. Review Supplementary Materials: Ensure that any supplementary files or materials provided with the submission also adhere to these anonymization guidelines.

8. File Properties Anonymization: To further enhance manuscript anonymity, authors should review and remove any metadata or file properties that may inadvertently reveal author information. This includes details such as document author names, comments, or other identifying information embedded within the document properties. Authors should use software tools or settings to inspect and eliminate such metadata before submitting the manuscript.

By meticulously following these guidelines, authors contribute to a fair and impartial peer review process, where manuscripts are evaluated solely based on their scientific content and merit. This anonymization process helps ensure that reviewers and editors can objectively assess the quality and validity of the research presented in the manuscript.

A cover letter providing information about the article should be prepared, including the article title, the name(s) of the author(s), title(s), affiliation(s), address, phone number, and email address, ORCID numbers, information about whether the article is based on a thesis (specialization, doctoral, etc.), and if applicable, where and when it was presented at a conference. Additionally, it should be specified who the corresponding author for correspondence is. Necessary information about the ethics committee approval of the research should also be provided.

Tables and figures should be presented in their original positions within the text. Figures will be published as they appear in the original. Therefore, original (first) copies of figures should be submitted. Figures in electronic versions of the accepted manuscripts may be in color; however, please note that the printed issues of the journal will be in black and white. Therefore, authors should keep this in mind when preparing and submitting their figures. It's important to ensure that figures are clear and convey information effectively in grayscale, as the color information may not be visible in the printed version. Authors are encouraged to use distinct patterns, shading, or labeling to differentiate elements within their figures to maintain their clarity and comprehensibility in both electronic and printed formats. This practice ensures that readers can fully grasp the content of the figures, regardless of whether they are viewing them electronically or in a physical print copy.

Examples of the citation format can be found at https://apastyle.apa.org/style-grammar-guidelines/references/examples.

For correspondence
Serkan Özel
Boğaziçi University, Faculty of Education, 34342 Bebek, Istanbul, Turkey
Tel: (+90 212) 359 45 88 Fax: (+90 212) 257 50 36
e-Mail: bujed@boun.edu.tr