Assessing second language speaking: Task type and language anxiety

Mehmet Gazan
National Defense University & Middle East Technical University, Türkiye

Vasfiye Geçkin
Izmir Democracy University, Türkiye

This study investigates the interaction between task type and foreign language anxiety (FLA) on speaking performance of learners of English as a foreign language (EFL). 70 pre-intermediate EFL learners (aged 17-21) enrolled in an English preparatory program at a state university in Türkiye completed the Foreign Language Classroom Anxiety Scale (FLCAS) and two speaking tasks: a monologic picture-cued description task comprising three prompts and a dialogic role-play task comprising three scenarios. The participants' oral performances were rated by three trained raters using task-specific analytic rubrics. A within-subjects design was employed, and the results of non-parametric tests of Wilcoxon signed-rank tests, Spearman's correlation, and Mann–Whitney U tests were reported. A moderate negative correlation was found between FLA and the monologic task performance while no significant relationship was observed between FLA and the performance on the dialogic task. Participants performed significantly better on the dialogic task although anxiety levels did not moderate this effect. These results suggest that task type may exert a stronger influence on speaking performance than language anxiety in structured assessment contexts. The findings offer practical implications for the design and sequencing of oral tasks in language testing and highlight the value of task format.

Keywords: task-based assessment, speaking performance, picture description, role-play, foreign language anxiety

Email address for correspondence: mehmet.gazan@metu.edu.tr

© The Author(s) 2026. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits the user to copy, distribute, and transmit the work provided that the original authors and source are credited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Introduction

Speaking is a fundamental component of any language curricula and is one of the focal and complex skills in language testing policies (Bachman & Palmer, 2010) since several variables tend to influence the second language (L2) speaking performance (Teng, 2007). Among these, linguistic factors such as transfer from first language (L1) (Dayat, 2017), affective factors such as foreign language anxiety (FLA) (Bielak, 2022), and pedagogical factors such as task type (Ko, 2023) can be listed. Tasks are defined as "devices for eliciting and evaluating communicative performances from learners in the context of language use that is meaning-focused and directed towards some specific goal" (Ellis, 2003, p. 279). Among a wide range of methods and formats employed as a part of testing procedures, tasks act as a crucial tool of assessment within the framework of task-based assessment (Norris, 2016). Tasks can be classified based on their goal orientations (i.e., convergent vs. divergent; Vo, 2018), purposes (i.e., open-ended vs. structured; Luoma, 2004), processing (i.e., personal, narrative and decision-making; Skehan & Foster, 1999), planning time (i.e., planned vs. unplanned task conditions; Li et al., 2015), the number of tested skills (i.e., independent vs. integrated; Brooks, 2009) and the number of test takers (i.e., monologic vs. dialogic; Galaczi, 2008; Sandlund et al., 2016).

When the number of test-takers is considered, the use of monologic tasks (e.g., story narration, oral presentation, and interviews) offers limited authenticity, dynamism, and interaction compared to real-life conversations (Brooks, 2009; Skehan, 2013), whereas paired/dyadic and group assessment tasks enable assessing larger samples of test-takers in a shorter time in real-life settings (Wigglesworth & Taylor, 2009), providing varied interaction samples (Galaczi, 2008) and more complex meaning negotiations among the L2 speakers (Sandlund et al., 2016). However, despite the flourishing research in monologic vs. dialogic assessment procedures, little attention has been given to comparing the two task-based assessment types. Therefore, the first aim of this study is to compare the role of task type on tertiary-level students' speaking performances in L2 English.

Second, FLA is reported to act either as a facilitative or a debilitative factor in testing speaking performance in an L2 (Oflaz, 2019; Tsang, 2022; Zhang, 2019). While anxiety

contributes to paying extra effort to complete a task, it may hamper the successful completion of the task (Shirvan & Talebzadeh, 2018). While much research has been dedicated to an investigation of FLA and L2 speaking assessment, further work is still needed to explore the interplay between different levels of anxiety and task type. Thus, the second aim of this study is to investigate the effect of anxiety on L2 speaking performance. The results are expected to shed light on the understanding of current testing and evaluation practices by providing practical insights into monologic and paired speaking tests.

Literature review

Task type and speaking performance

As a result of the prevalence of communicative and real-life oriented foreign language teaching and testing methods, various foreign language speaking task types such as picture descriptions, role-plays, group discussions, and read-aloud tasks have been a part of task-based assessment (Pawlak, 2016). Yet there appears to be incongruence in findings regarding the impact of task type on L2 oral performance across different speaking tasks.

Some studies have reported the existence of such an impact (e.g., Bottini, 2022; In'nami & Koizumi, 2016; Ko, 2023) while other studies (e.g., Ahangari & Semiyari, 2019; Khoram, 2019; Susoy, 2018; Tavakoli et al. 2017) have not found a significant difference across various tasks. First, looking at those studies that demonstrated a task effect, Ko (2023) found that Korean learners (n = 57) of EFL performed better in a dialogic task than in a monologic task in terms of topic development, fluency, and accuracy. Similarly, Teng (2007) reported significant task-related differences in analytic measures of speaking performance, with L2 learners (n = 30) achieving higher fluency and complexity scores in answering questions tasks than in presentation and picture description tasks. Bottini (2022) reported that L2 speakers exhibited less diversity and sophisticated vocabulary in the dialogic task than in the monologic one. In addition, Mathews and Sudharshana (2021) investigated how task type and the interaction between monologic and dialogic tasks affected the speech production of Indian EFL learners (n = 56). The monologic narrative task productions are reported to be more complex but less accurate when compared to the dialogic ones. Previous

work also suggests that level of L2 proficiency is a predictor of speaking performance with respect to different task types. For instance, the speaking skills of EFL learners (n = 30) at three different levels of proficiency were tested on three different task types which required expressing personal information, giving a narrative and making a decision (Garcia-Ponce & Tavakoli, 2022). The learners performed the personal information task with the highest accuracy and fluency while they completed the narrative task with lower accuracy and fluency yet with greater syntactic diversity. Overall, the more proficient the learners got, the more fluently and accurately they performed.

On the other hand, some research pointed to no significant effect of task type on speaking performance. To illustrate, Tavakoli et al. (2017) examined the speaking performance of L2 learners (n = 32) who took four monologic tasks with increasing difficulty. Despite noticeable differences in speed, breakdown, and repair measures, the speed of speech, length of pauses, and frequency of pauses remained unaffected by the task type. Ahangari and Semiyari (2019) tested the speaking performance of Iranian EFL learners (n = 40) across tasks which aimed at problem-solving, picture-describing, explaining, and storytelling. Neither the rubric type nor the task type had a meaningful effect on L2 speaking performance, except for story narration.

FLA and speaking performance

No consensus has been reached about the role of foreign language anxiety (FLA) on L2 speaking performance despite extensive research within the past few decades. However, numerous studies have demonstrated negative effects. Zhang (2019) found an overall negative correlation between FLA and foreign language performance irrespective of language proficiency levels. Similarly, Oflaz (2019) reports a moderate but significant negative correlation between Turkish EFL learners' (n = 110) anxiety levels. Damayanti and Listyani (2020) pointed out that communication apprehension, test anxiety and fear of negative evaluation triggered Indian EFL students' (n = 52) foreign language anxiety levels. In another study, Bielak (2022) purported the negative effects of anxiety on speaking performance after analyzing fluency measures such as length of pauses, mean length of run and rate of articulation.

The findings of negative correlations between FLA and fluency measures provided support for the idea that FLA may hinder L2 speaking performance. In a more elaborate recent study, Rood and de Jong (2023) explored the interplay between FLA and fluency levels of Dutch EFL learners (n = 22) through two similar tasks that required a comparison of two pictures and scene-building in which each participant had distinct pictures to create a narrative that ultimately resulted in an identical outcome. They were designed in such a way that task complexity would not interfere with speaking performance. The manipulation of FLA affected fluency which in turn led to increased filled pauses. What is more, higher levels of anxiety were correlated with a greater utilization of silent pauses. The simulated recalls in the high-anxiety condition revealed a greater number of problems concerning the message's substance compared to those in the low-anxiety condition. Lastly, Tsang (2022) addressed the relationship between EFL learners' anxiety in public speaking, their competence in self-perceived pronunciation skills and actual speaking performance. Data coming from questionnaires and interviews showed that self-perceived pronunciation skills and anxiety were negatively correlated.

On the other hand, only a few studies have concluded that anxiety acted as a facilitative factor on L2 speaking performance (Shazly, 2021; Zhang & Liu, 2013). According to Eysenck (1979), learners put more effort into tasks when they develop an awareness of the extent anxiety affects their performance and that anxiety could potentially improve their performance. Zhang and Liu (2013) explored the link between oral test anxiety and speaking strategies of Chinese university students (n = 1697). Results revealed that lower proficiency level students were more adversely affected by FLA compared to their higher proficiency peers. Finally, Shazly (2021) examined the speaking performance of EFL learners (n = 48) who got engaged in conversations with AI chatbots through role-plays. The findings revealed positive effects of anxiety on L2 oral performance in enhancing students' cognitive abilities and linguistic skills.

Concerning the interplay between task type and FLA, Brennan (2016) tested college students (n = 108) who were given problem-solving tasks of varying complexity. The more complicated the task got, the higher the levels of anxiety that were observed. In another study, Adnyani et al. (2014) examined the role of scripted role-play tasks and FLA on the speaking competence of EFL learners (n = 68). There existed a significant

interaction between the role-play task and FLA, which particularly benefited students with high anxiety levels. Song and Lee (2015) added another variable to the research by investigating the interaction between task type and task practice on L2 oral production skills of Korean EFL learners (n = 32) whose speaking proficiency was assessed through a narration and a description task. While task practice did not substantially enhance overall speaking performance, anxiety levels dropped thanks to intensive practice on task.

Overall, whereas some studies suggest a positive contribution of task type to speaking performance, others indicate that task-related differences may not solely determine speaking performance.

One further consideration is the specific educational context in which this study was conducted. A military school represents a unique learning environment shaped by hierarchies, uniform behavioural expectations, and a dual focus on academic and military performance. Despite these distinct features, research on FLA in military settings remains limited. Ivanova et al. (2022), for example, investigated EFL learners at a military academy in North Macedonia and found that learners experienced high anxiety due to fear of negative evaluation, lack of confidence, and limited communicative practice. Similarly, Alqahtani (2018) reported that Saudi military cadets were influenced by institutional demands and extrinsic pressure, which contributed to heightened anxiety levels. Complementing these findings, Sipos et al. (2019) studied U.S. soldiers and found elevated levels of burnout and anxiety, attributing them to the academic rigor and military constraints imposed on learners. These studies highlight the importance of accounting for institutional context when researching FLA. Given the scarcity of research in this area, the present study contributes to the literature by examining anxiety and speaking performance in a military preparatory school in Türkiye—a setting where discipline, conformity, and institutional expectations can shape the language learning experience.

To further explore the role of FLA in the oral performance of L2 learners across various task types, the present study examines the relationship between task type and FLA among pre-intermediate level EFL learners in the Turkish context. The aim is to investigate the extent task type and FLA interact by subjecting all the learners to the same testing procedure. To this end, the following research questions were formulated:

(1) Which components of FLA are associated with pre-intermediate EFL students' speaking performance at a military school?

(2) How is FLA related to students' speaking performance across monologic and dialogic tasks?

(3) To what extent does the interaction between the task type and anxiety impact the oral performance of the learners?

Method

In this study, the task type and FLA scores were taken as the predictor variables and the speaking scores assigned to the participants were taken as the response variable. The study utilizes a *within-subjects design* with repeated measures, where the same learners were exposed to each predictor variable (i.e., two different speaking tasks) multiple times. Although the study is primarily quantitative, brief follow-up interviews with six participants were also conducted to gain further insights.

Participants

Data was collected from 70 pre-intermediate EFL learners enrolled in an English preparatory program at a military school in Türkiye. Initially, 40 students were selected through convenience sampling and all identified Turkish as their native language. Following reviewer feedback on sample size limitations, an additional 30 participants were recruited using the same procedures to ensure consistency. The group consisted of male and female EFL students who were enrolled in an intensive foreign language program that provided 27 hours of English instruction per week over eight months. Participants' ages ranged from 17 to 21 (M = 19.4, SD = 2.3). Based on institutional assessments, the average English proficiency score of the participants was 82.83, with a range of 70 to 93, which the institution classifies as B1 according to its CEFR-based criteria. On average, the participants reported their first exposure to English at age 11.63 (SD = 1.8, range= 6 -15), and reaching a functional level of fluency around an average 18.01 years of age (SD = 0.94, range = 14 - 21). The learners were also asked about their language use habits. They reported using English for 0 to 90 percent (M = 23.41%, SD = 21.50) and Turkish for 10 to 100 percent (M = 76.37%, SD

= 22.50) in their daily communication. The students self-rated their level of EFL proficiency at level 2 (i.e., basic, $N = 3$), at level 3 (i.e., intermediate, $N = 28$), at level 4 (i.e., upper-intermediate, $N = 27$), and at level 5 (i.e., advanced, $N = 12$). This demographic profile demonstrates that the learners had a homogeneous background in terms of their language proficiency and daily language use patterns.

Instrument

The Foreign Language Classroom Anxiety Scale (FLCAS) was first developed by Horwitz et al. (1986) and later modified by Na (2007). It is used to measure learners' anxiety across four dimensions: communication apprehension (items 1, 9, 14, 18, 24, 27, 29, 32), fear of negative evaluation (items 3, 7, 13, 15, 20, 23, 25, 31, 33), test anxiety (items 2, 8, 10, 19, 21), and in-class anxiety (items 4, 5, 6, 11, 12, 16, 17, 22, 26, 28, 30). The instrument consists of 33 items rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Eight negatively worded items (Items 5, 8, 11, 14, 18, 22, 28, and 32) were reverse coded prior to the analysis so that higher values consistently reflected higher anxiety. Following Marcos-Llinás and Juan-Garau's (2009) classification, FLCAS total scores were also used to categorize participants into three anxiety groups: scores between 33-77 indicated low anxiety, those between 78-121 moderate anxiety, and those between 122-165 high anxiety. These groupings allowed for categorical interpretation of overall anxiety levels in addition to continuous score analysis. Rather than relying solely on total scores, the data was also analysed at the subscale level using scale means and standard deviations to enhance interpretability. Each subscale score was calculated by averaging the responses to the corresponding items. The descriptive statistics and internal consistency reliability estimates (Cronbach's alpha) for each subcomponent were as follows: communication apprehension ($M = 2.69$, $SD = 0.66$, $\alpha = .76$), fear of negative evaluation ($M = 2.35$, $SD = 0.63$, $\alpha = .71$), test anxiety ($M = 2.49$, $SD = 0.41$, $\alpha = .72$), and in-class anxiety ($M = 2.47$, $SD = 0.48$, $\alpha = .68$). All subscales demonstrated acceptable internal consistency and were used in subsequent correlation and comparative analyses.

Tasks

In line with the research questions, a monologic and a dialogic task were used to tap interactional, discourse and pragmatic competence in L2 speech. Table 1 summarizes the task types and characteristics.

Table 1. Tasks characteristics (Adapted from Kley, 2015)

Type	Requirements	Goal orientation	Task orientation	Format
Picture-cued description	Describing the two pictures & comparing and contrasting them	Divergent	Monologic	Individual
Role-play	Following the steps written in the prompt cards and acting out the dialogues in dyads	Convergent	Dialogic	Paired

As shown in Table 1, the picture-cued description task was divergent, monologic and individualistic. Divergent tasks, as described by Hommel (2012) and Swan (2005), necessitate acquiring novel significant knowledge and offer multiple outcome possibilities, allowing for diverse cognitive styles and potentially yielding varied results. On the other hand, the role-play task was convergent and dialogic (i.e., completion of the task necessitated interaction between the test-takers). Namely, the tasks typically have only one correct answer and involve collaborative work with relatively simple responses that do not demand high cognitive effort or extensive referencing. The aim was to achieve consensus among learners to produce a reasonable solution (Hommel, 2011).

Picture-cued description task

The task had three steps and comprised six sets of pictures which were used to make comparisons such as eating at a fancy restaurant vs. at a fast-food restaurant and shopping online vs. shopping at a mall (see Appendix A). Test items, testing procedures and scoring rubrics were developed in accordance with the expert opinions in the field. The task allowed the test takers to visualize concepts in an open-ended manner, offering visibility and generating interest in practicing the foreign language (Lavalle & Briesmaster, 2017). It helps set up the context with sufficient background information. The selection of pictures from actual testing materials enhanced the authenticity and validity of the assessment instrument.

Role-play task

The second type of task involved semi-structured paired role-plays and promoted interaction in a dyadic format (see Appendix B). Similar to in-class materials, three role-play situations were implemented: buying clothes, ordering food, and inviting a friend to an event. The aim of choosing this task was to contrast the monologic nature of the first data collection tool and highlight task-related score differences (if any). Role plays stimulate participants to produce spontaneous speech in accordance with the given roles and serve communicative functions such as persuasion and complaints (Shohamy et al., 1986). The task did not require scripts, which yielded authentic and smooth performances in real-life contexts (Toding et al., 2021).

A pilot study of the two tasks was conducted with 10 EFL students who were enrolled in the same English preparatory program, but who did not participate in the main study. The pilot task performances were evaluated by two experienced EFL instructors. Based on the pilot results, no changes were made to the tasks as they were found to be clear and appropriate.

Scoring

Rubrics

An analytic rubric was utilized to provide a comprehensive range of scoring criteria for assessing learner performances across tasks. In alignment with Alderson and Banerjee's (2002) argument that different task types necessitate distinct assessment criteria, two task-specific yet structurally similar rating scales were employed. These were adapted from Torky's (2006) oral performance rubric. The rubric used for the picture-cued description task evaluated grammatical competence, discourse competence, pragmatic competence, fluency, and task achievement. The rubric for the role-play task included these same components, with the addition of interactive communication to reflect the dialogic nature of the task. To ensure content validity, the adapted rubrics were reviewed by a panel of seven ELT professionals with expertise in speaking assessment. Based on their feedback, several refinements were made to enhance the comprehensiveness and clarity of the scoring descriptors (see Appendix C).

Raters

Three independent Turkish-English bilingual EFL teachers who had previously engaged in speaking exams at the higher educational institutions were recruited as raters. They ranged between 27 to 29 years of age and had a minimum of 5 years of teaching experience at college level. These three raters held at least an M.A. degree from an English Language Teaching Program. The raters were given a standardization session to ensure familiarity with the rubric and homogeneity in scoring. The student scores obtained in the pilot study were employed in the standardization session, which was conducted online and lasted for an hour. During this session, the raters jointly scored sample performances, discussed discrepancies, and calibrated their interpretations of the rubric criteria to ensure alignment in scoring practices. This training aimed to minimize subjectivity and promote consistency across raters before evaluating the main dataset. Inter-rater reliability analyses using Intraclass Correlation Coefficient (ICC) were later conducted to verify consistency. It was based on a two-way mixed-effects model with absolute agreement. Since three independent raters evaluated all the participants, the average measures ICC was calculated. The overall ICC for the averaged scores across raters was .71, which is interpreted as indicating moderate agreement (Shrout & Fleiss, 1979). The 95% confidence interval for this value ranged from .44 to .85. An inspection of individual rater tendencies revealed that Rater 2 awarded slightly lower scores (M = 78.05, SD = 10.62) than Rater 1 (M = 84.25, SD = 10.80) and Rater 3 (M = 86.67, SD = 6.03), contributing to lower pairwise agreement between Rater 2 and the others.

Procedure

The original data was collected by the first author in two consecutive days on campus in a quiet room allocated to the researcher. After taking an institutional placement test, 130 students completed the FLCAS. Out of these 130 students, 50 volunteered to take part in the speaking test. However, due to incomplete recordings and missing questionnaire responses, only 40 students were included in the initial round of analysis. Following reviewer feedback regarding the limited sample size, an additional 30 participants were recruited using the same procedures and institutional setting, resulting in a final sample of 70 participants. For the dialogic speaking task, learner pairings were formed based on participants' preferences. Ten pairs were tested each

day. In the monologic task, the first author acted as the interlocutor by presenting the pictures and asking follow-up questions. In the subsequent role-play task, the first author provided the task instructions and role cards and monitored the interaction, while the paired students acted as interlocutors for each other. The testing sessions were audio recorded.

The testing session started with a warm-up question on topics like evening routines or preferences in books and films. After the instructions were reviewed, one-minute preparation time was given for each participant prior to the monologic picture-cued description task. The participants were given six sets of pictures, without the option to take notes during the pre-speaking planning phase. The task required longer turns to compare and describe parallel images within a specific time limit of around one minute. As one participant delivered an extended turn, the other member of the pair awaited their turn without interruption. The interlocutor posed a question related to the pictures, granting approximately 30 seconds of speaking time to the other participant. Roles were subsequently swapped, with the first participant answering questions related to the 1st, 3rd, and 5th picture prompt, while the other participant completed the 2nd, 4th, and 6th prompts. Only the extended turns were evaluated in this process.

Following the picture-cued description task, the same pairs engaged in self-selected dyadic role-play scenarios. Task expectations were outlined, roles were assigned through role cards featuring prompts, and two minutes were granted for preparation and potential inquiries. Once ready, the participants enacted their roles, with their performances remaining uninterrupted across three role-play scenarios in this part. The paired role-play task performances were free from any time constraints.

The three raters assessed the students' performance on the two tasks based on the audio-recordings and then the scores were entered anonymously into Excel sheets for the statistical analyses.

Data analysis

To assess whether the data met assumptions of normality, Shapiro–Wilk tests were conducted on each rater's average task scores. The results revealed that several

distributions, particularly for the role-play scores from Rater 1 (W = .946, p = .005) and Rater 3 (W = .941, p = .003), significantly deviated from normality. Given the mixed results and to follow a conservative approach, non-parametric tests were used in the main analyses to account for potential violations of normality.

To address the first research question, the sum of scores within each category of the FLCAS was calculated, and the total scores were divided by the number of items in each category. As for the second research question, Spearman's rank-order correlation was used to assess the relationship between students' FLCAS scores and their speaking task performance, given the non-parametric nature of the data. For the analysis of the third research question, Wilcoxon signed-rank tests were used to determine if there were statistically significant differences in students' scores as evaluated by each rater. Finally, Mann Whitney-U tests were conducted to explore the relationship between task types and anxiety groups.

Results

In this section, the findings related to each research question are reported.

RQ1: Which components of FLA are associated with pre-intermediate EFL students' speaking performance at a military school?

The participants in this study had an overall moderate level of FLA (M = 82.23, SD = 14.98). Out of 70 students, 23 participants (32.9 %) were classified as having low anxiety (M = 64.78, SD = 8.98), and 47 participants (67.1 %) exhibited moderate anxiety levels (M = 90.77, SD = 8.45). No students in the sample were categorized as having high anxiety. Student scores were further analysed based on the subcomponents of the FLCAS: communication apprehension, test anxiety, fear of negative evaluation, and in-class anxiety. These components were used to explore which dimension of FLA most strongly impacted learners' speaking performance in subsequent analyses. Since the distribution of student scores across components violated the assumptions of normality, Table 2 gives a descriptive summary of students' levels of FLA across four components.

Table 2. Breakdown of student responses to FLCAS

Component	Number of Items	Range	Mean (*SD*)	Overall Mean (*SD*)
Communication apprehension	8	10-33 (out of 40)	2.69 (0.66)	21.50 (5.24)
Test anxiety	5	9-19 (out of 25)	2.35 (0.63)	12.46 (2.05)
Fear of negative evaluation	9	9-34 (out of 45)	2.49 (0.41)	21.13 (5.63)
In-class anxiety	11	15-43 (out of 55)	2.47 (0.48)	27.14 (5.30)

Note. Mean (SD) refers to the average item score per participant for each component. Overall Mean (SD) refers to the total component score per participant, based on all items in that component.

As shown in Table 2, the anxiety scores across the four dimensions were moderately distributed, with mean subscale scores ranging from approximately 12 to 27. Students reported the highest anxiety in in-class contexts (M = 27.14, SD = 5.30), followed by communication apprehension (M = 21.50, SD = 5.24). The dimension of fear of negative evaluation ranked third (M = 21.13, SD = 5.63), while test anxiety elicited the lowest mean scores (M = 12.46, SD = 2.05). These results suggest that classroom participation and situations which require spontaneous communication were perceived as more anxiety-inducing than evaluative contexts like testing.

When anxiety scores were analyzed based on overall speaking performance, some variation emerged across the three proficiency tiers. Students in the high-performance group reported lower levels of communication apprehension (M = 20.48) compared to both the medium (M = 21.91) and low performers (M = 22.13). Interestingly, in-class anxiety was highest among the average performers (M = 28.96), while test anxiety remained relatively stable across all groups, ranging narrowly from M = 12.39 to M = 12.52. The dimension of fear of negative evaluation also showed minimal variation, with slightly higher scores in both the low and high-performing groups.

RQ2: How is FLA related to students' speaking performance across monologic and dialogic tasks?

A Spearman correlation analysis was conducted to examine the relationship between students' FLCAS scores and their speaking performance across the two task types[3]. The results revealed a small, statistically significant negative correlation between FLCAS scores and the monologic picture-cued description task (ρ = –.267, CI [–.47, –.03], p

[3] Confidence intervals were calculated using bootstrapped estimates to align with current estimation-focused reporting practices (Cumming & Calin-Jageman, 2024).

= .026), indicating that students with higher levels of anxiety tended to perform more poorly on this task type. In contrast, the correlation between FLCAS scores and the dialogic role-play task was weaker and did not reach statistical significance (ρ = –.203, CI [–.47, –.03], p = .094). These findings suggest that anxiety may have a more pronounced negative impact on monologic speaking tasks, which require extended turns and less interactional support compared to dialogic role-play tasks.

Table 3. The correlation between FLCAS and the task scores

		Picture Description Task			Role-play Task		
		Prompt 1	Prompt 2	Prompt 3	Scenario 1	Scenario 2	Scenario 3
FLCAS scores	ρ	.248	-.285	-.194	.093	-.236	-.225
	p	.040	.018	.109	.447	.051	.063

Note: Spearman's ρ values represent effect sizes. Exact p-values are reported; statistical significance should be interpreted in relation to effect magnitude and study context.

As shown in Table 3, all task scores were negatively associated with students' FLA levels except for the first prompt and first scenario, which showed weak positive relationships. For the picture-cued description task, the second prompt showed the strongest negative correlation (ρ = –.285, 95% CI [–.49, –.05], p = .018), followed by the third (ρ = –.194, 95% CI [–.41, .04], p = .109). The first prompt, in contrast, exhibited a small positive correlation (ρ = .248, 95% CI [.01, .46], p = .040), indicating that higher anxiety was unexpectedly associated with slightly better performance on that prompt. For the role-play task, the second (ρ = –.236, 95% CI [–.45, .00], p = .051) and third (ρ = –.225, 95% CI [–.44, .01], p = .063) scenarios demonstrated modest negative associations with anxiety, whereas the first scenario showed a very weak positive correlation (ρ = .093, 95% CI [–.15, .32], p = .447). Overall, these findings suggest that the negative impact of FLA was more consistent and stronger across later steps, particularly within the picture-cued format. This pattern may reflect increasing cognitive load or diminished procedural benefits over time, especially in monologic contexts where extended individual performance is required.

RQ3: To what extent does the interaction between the task type and anxiety impact the oral performance of the learners?

To compare the students' speaking performance obtained from the repeated measures of both tasks, Wilcoxon signed-rank tests were conducted. The statistical output obtained from these tests is summarized in Table 4.

Table 4. The interaction between task type and L2 speaking performance

Raters	Median (Mean ranks) Scores (SD)			
	Picture Description	Role-play	*Z*	*p*
Rater 1	80.67 (10.96)	83.33 (10.56)	3.584	< .001
Rater 2	76.00 (10.73)	80.00 (8.86)	4.277	< .001
Rater 3	82.00 (8.28)	84.67 (8.40)	1.741	.082

Note. Z values reflect Wilcoxon signed-rank tests comparing performance across tasks. Effect sizes are reported as rank-biserial correlations (r). Interpretations are based on estimation thinking (Cumming & Calin-Jageman, 2024) with confidence intervals provided.

Results of the Wilcoxon signed-rank tests showed that students performed significantly better in the role-play task according to Rater 1 ($Z = 3.584$, $p < .001$, $r = .428$, 95% CI [0.21, 0.60]) and Rater 2 ($Z = 4.277$, $p < .001$, $r = .511$, 95% CI [0.31, 0.67]), with moderate to large effect sizes. For Rater 3, the difference was not statistically significant ($Z = 1.741$, $p = .082$), though a small effect size was observed ($r = .208$, 95% CI [−0.03, 0.42]). These findings indicate that learners generally performed better in dialogic, interactive task items than in the monologic ones, and that the effect of task type on performance was consistent across most raters. We also examined whether task type and anxiety interacted to affect pre-intermediate EFL learners' L2 speaking performance, using a Mann-Whitney U test (see Table 5).

Table 5. The interaction between task Type and anxiety

Task type	Anxiety level	*N*	Mean ranks	Sum of ranks	SD	*p*
Picture description	Low	23	33.75	830.5	12.54	.440
	Moderate	47	33.71	1584.5	10.23	
Role-play	Low	23	38.20	840.5	10.18	.360
	Moderate	47	33.50	1574.5	10.75	

Table 5 illustrates that students' speaking performance did not significantly differ across anxiety levels in either task type. For the picture-cued description task, the low anxiety group ($N = 23$, M rank = 37.75, $SD = 12.54$) and the moderate anxiety group ($N = 47$, M rank = 33.71, $SD = 10.23$) did not demonstrate a statistically significant difference ($p = .440$), with a small effect size ($r = −0.117$, 95% CI [−0.351, 0.117]). Similarly, in the role-play task, although the low anxiety group (M rank = 38.20, $SD = 10.18$) slightly outperformed the moderate anxiety group (M rank = 33.50, $SD = 10.75$),

this difference was also not statistically significant (p = .360), with a small effect size (r = −0.136, 95% CI [−0.370, 0.098]).These findings suggest that a mild tendency toward better performance among learners with lower anxiety. However, given the small effect sizes and lack of significance, these observations should be interpreted cautiously.

To enrich the interpretation of these findings, an online focus group interview session was conducted with six participants. When briefly asked what made them anxious in L2 speaking, and whether they felt more comfortable when speaking alone or with a peer, most reported greater comfort and fluency in paired tasks. These perceptions are in line with the slight performance advantage observed in the role-play condition.

Discussion

This study aimed to examine whether task type, FLA, and their interaction influenced students' speaking performance. Using a within-subjects repeated measures design, all the participants completed both monologic and dialogic speaking tasks. The most notable contribution of this research lies in its focus on FLA within a military school setting—an underexplored context in second language acquisition—which involved a relatively homogeneous sample of 70 pre-intermediate learners with low to moderate anxiety levels. These findings align with earlier studies conducted in Türkiye (e.g., Çakıcı, 2016; Elaldi, 2016), possibly due to shared institutional expectations, similar educational policies, and consistent language-learning experiences among Turkish EFL learners. The analysis of the FLCAS subcomponents revealed that communication apprehension was the most anxiety-inducing dimension, consistent with studies such as Aeni et al. (2017) and Noor et al. (2015). This pattern may reflect the limited opportunities for communicative language use in Turkish higher education, where grammar-focused instruction is still dominant. In contrast to findings by Na (2007) and Shabani (2012), fear of negative evaluation ranked lower in this study. This might be attributed to the military school's structured evaluation system, consistent grading practices, and the confidence-building nature of institutional culture.

Although moderate levels of anxiety are sometimes linked to improved performance (e.g., Zhang & Liu, 2013; Marcos-Llinás & Juan-Garau, 2009), the present study found no significant performance differences between low- and moderate-anxiety groups in

either task type. While initial picture prompt and role-play scenario showed weak facilitative effects, possibly due to the effects of procedural repetition (Ellis, 2009), this benefit diminished in subsequent tasks. This could be likely as a result of increased cognitive demands and topic novelty, echoing Hu's (2018) observation that fluency gains are not guaranteed across task repetitions. The significant score differences between task formats support earlier research highlighting the impact of task type on speaking performance (Bottini, 2022; Garcia & Ponce, 2022; Ko, 2023). In line with Ko (2023) and Mathews and Sudharshana (2021), the participants performed better on dialogic task items, likely due to their interactive, authentic nature. While monologic tasks have been favored in some contexts (Martin, 2019; Phuong, 2018), the paired nature of role-plays in this study may have allowed learners to stay within their comfort zones by interacting with peers rather than solely with the test administrator. Hence, this study reinforces the importance of task type in L2 assessment. Although task complexity is known to influence both performance and anxiety (Brennan, 2016), the more complex dialogic task in this study yielded better performance without significant anxiety-based differences. This may be attributed to the discipline, self-regulation, and uniform preparation of military students, whose consistent language habits may have buffered anxiety's influence across task types. Overall, the findings suggest that, while FLA and task type interact in complex ways, task format may exert a more robust effect on speaking performance in structured assessment settings.

Conclusion

This study set out to examine the effects of FLA, task type, and their interaction on the speaking performance of pre-intermediate EFL learners. It addressed the broader question of how affective and task-related factors together shape L2 speaking performance in classroom assessment contexts. The results showed that while learners' anxiety levels were relatively low, task type played a more decisive role in influencing their speaking performance. From a pedagogical perspective, our findings suggest that speaking assessments should incorporate both dialogic and monologic tasks to capture different dimensions of communicative ability. Dialogic tasks, in particular, may promote greater engagement and authenticity by allowing lower-proficiency learners to perform more confidently. These insights also remind us that assessment tasks are not merely tools for measurement but opportunities that can either heighten or

alleviate learners' anxiety, shaping how they experience language use. This study is not without limitations. The participant group was relatively small and drawn from one institutional context, which may limit the generalizability of the findings. Future research could employ mixed-method or longitudinal designs to explore how FLA, task engagement, and feedback interact across time and varying proficiency levels. In conclusion, this study contributes to a growing understanding that affective factors and task design jointly influence language performance. Recognizing this interplay can help educators and test developers design speaking assessments that are both valid and supportive of learners' communicative development.

Acknowledgements

We extend our gratitude to the raters and participant students.

Author disclosures

The authors have no conflicts of interest to disclose.

The necessary ethics permissions were received from (i) Izmir Democracy University, dated 01.04.2021 and numbered 2021/27 and (ii) the university where the study was conducted (dated 29.12.2020 and numbered 25592990-604.02-E.778993).

The study was supported by 2210-A TUBITAK merit-based scholarship program for M.A. students and is based on the first author's M.A. thesis titled "The Task-Based Assessment of Speaking Skills: Do Anxiety and Task Type Matter?", supervised by the second author. The authors contributed as follows:

1st author: writing & editing original draft, data collection, data analysis, funding acquisition. 2nd author: supervision, project administration, review & editing, conceptualization.

ORCID iDs

Mehmet Gazan https://orcid.org/0000-0002-7233-9033
Vasfiye Geçkin https://orcid.org/0000-0001-8532-8627

References

Adnyani, N. W., Seken, M., & Nitiasih, M. (2014). The effect of scripted role-play and foreign language anxiety on the eleventh-grade students' speaking competence at sma n 8 denpasar in academic year 2013/2014. *Jurnal Pendıdıkan Dan Pembelajaran Bahasa Indonesia, 3*(1), 1-11.

Aeni, N., Jabu, B., Rahman, M. A., & Strid, J. E. (2017). English oral communication apprehension in students of Indonesian maritime. *International Journal of English Linguistics, 7*(4), 158. https://doi.org/10.5539/ijel.v7n4p158.

Ahangari, S., & Semiyari, S. R. (2019). The impact of task types and rating methods on Iranian EFL learners' speaking scores. *Research in English Language Pedagogy, 7*(2), 187-208. https://doi.org/10.30486/relp.2019.665887

Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching, 35*(2), 79-113. https://doi.org/10.1017/S0261444802001751

Alqahtani, A. F. (2018). *English language learning motivation and English language learning anxiety in Saudi military cadets: A structural equation modelling approach* [Unpublished doctoral dissertation]. University of Southampton.

Bachman, L.F., & Palmer, A.S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world.* Oxford University Press.

Bielak, J. (2022). To what extent are foreign language anxiety and foreign language enjoyment related to L2 fluency? An investigation of task-specific emotions and breakdown and speed fluency in an oral task. *Language Teaching Research, 29*(3), 911-941. https://doi.org/10.1177/13621688221079319

Bottini, R. (2022). *Lexical complexity in L2 English speech: Evidence from the trinity Lancaster corpus*. [Unpublished doctoral dissertation]. Lancaster University.

Brennan, K. A. (2016). *Second language anxiety and task complexity*. [Unpublished doctoral dissertation]. Universitat de Barcelona.

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing, 26*(3), 341-366. https://doi.org/10.1177/0265532209104666

Çakıcı, D. (2016). The correlation among EFL learners' test anxiety, foreign language anxiety and language achievement. *English Language Teaching, 9*(8), 190-203. https://doi.org/10.5539/elt.v9n8p190

Cumming, G., & Calin-Jageman, R. (2024). *Introduction to the new statistics: Estimation, open science, and beyond*. Routledge. https://doi.org/10.4324/9781032689470

Damayanti, M. E., & Listyani, L. (2020). An analysis of students' speaking anxiety in academic speaking class. *ELTR Journal*, *4*(2), 152-170. https://doi.org/10.37147/eltr.v4i2.70

Dayat, D. (2017). Analysis on English speaking performance: Exploring students' errors and the causes. *Journal of Education, Teaching and Learning, 2*(1), 71-74.

Elaldi, S. (2016). Foreign language anxiety of students studying English language and literature: A sample from Turkey. *Educational Research and Reviews, 11*(6), 219-228. https://doi.org/10.5897/ERR2015.2507.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474-509. https://doi.org/10.1093/applin/amp042

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

Galaczi, E. D. (2008). Peer–peer interaction in a speaking test: The case of the first certificate in English examination. *Language Assessment Quarterly, 5*(2), 89-119. https://doi.org/10.1080/15434300801934702

Garcia-Ponce, E. E., & Tavakoli, P. (2022). Effects of task type and language proficiency on dialogic performance and task engagement. *System*, *105*, Article 102734. https://doi.org/10.1016/j.system.2022.102734

Hommel, B. (2012). Convergent and divergent operations in cognitive search. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 56, pp. 221–256). Academic Press.

Horwitz, E. K., Horwitz, M. B., & Cope, I. (1986). Foreign language classroom anxiety. *Modern Language Journal*. 70, 125-32. https://doi.org/10.2307/327317

Hu, X. (2018). *Language learning through task repetition: Effects of task type, task-type repetition, and performance criteria on L2 oral production*. John Benjamins.

In'nami, Y., & Koizumi, R. (2016). Task and rater effects in L2 speaking and writing: A synthesis of generalizability studies. *Language Testing, 33*(3), 341-366. https://doi.org/10.1177/0265532215587390

Ivanova, B., Trajanovska, E., & Kirova, S. (2022). Investigation of students' motivation and anxiety at the Military Academy 'General Mihailo Apostolski' while learning English as a foreign language. *Vospitanie/Vospitanie - Journal of Educational Sciences, Theory and Practice*, *15*(21), 51-64. https://doi.org/10.46763/JESPT22172092i

Khoram, A. (2019). The impact of task type and pre-task planning condition on the accuracy of intermediate EFL learners' oral performance. *Cogent Education*, *6,* 1-13. https://doi.org/10.1080/2331186X.2019.1675466

Kley, K. (2015). *Interactional competence in paired speaking tests: Role of paired task and test-taker speaking ability in co-constructed discourse* [Unpublished doctoral dissertation]. University of Iowa.

Ko, H. (2023). English oral proficiency measured by holistic and analytic assessments in dialogic and monologic tasks. *English Teaching, 78*(1), 63-82. https://doi.org/10.15858/engtea.78.1.202303.63

Lavalle, P. I., & Briesmaster, M. (2017). The study of the use of picture descriptions in enhancing communication skills among the 8th- grade students—learners of English as a foreign language. *Inquiry in Education, 9*(1), 1-20.

Li, L., Chen, J., & Sun, L. (2015). The effects of different lengths of pre-task planning time on L2 learners' oral test performance. *TESOL Quarterly, 49*(1), 38-66. https://doi.org/10.1002/tesq.159

Luoma, S. (2004). *Assessing speaking*. Cambridge. https://doi.org/10.1017/cbo9780511733017

Marcos-Llinás, M., & Juan-Garau, M. (2009). Effects of language anxiety on three proficiency-level courses of Spanish as a foreign language. *Foreign Language Annals, 42*(1), 94-111. https://doi.org/10.1111/j.1944-9720.2009.01010.x

Martin, E. A. (2019). *Using culturally responsive visual aids to increase English learners' speaking assessment scores.* [Unpublished master's thesis]. California State University.

Mathews, S. M., & Sudharshana, N. P. (2021). Effect of task structure and interaction conditions on oral performance. In N. P. Sudharshana, & L. Mukhopadhyay (Eds.), *Task-based language teaching and assessment*. (pp. 275-306). Springer. https://doi.org/10.1007/978-981-16-4226-5_14

Na, Z. (2007). A study of high school students' English learning anxiety. *The Asian EFL Journal, 9(*3), 22-34.

Noor, A., Rafek, M., Khalid, P. Z., & Mohammad, R. (2015). Communication apprehension in language learning: Is it serious? *International Journal of Arts and Sciences, 08*(02), 311-320.

Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics, 36,* 230-244. https://doi.org/10.1017/S0267190516000027

Oflaz, A. (2019). The effects of anxiety, shyness and language learning strategies on speaking skills and academic achievement. *European Journal of Educational Research*, *8*(4), 999-1011. https://doi.org/10.12973/eu-jer.8.4.999

Pawlak, M. (2016). Assessment of language learners' spoken texts: Overview of key issues. In H. Chodkiewicz, P. Steinbrich, & M. Krzemińska-Adamek (Eds.), *Working with text and around text in foreign language environments*. (pp. 89-105). Springer. https://doi.org/10.1007/978-3-319-33272-7_6

Phuong, B. H. (2018). Can using picture description in speaking sessions help improve EFL students' coherence in speaking? *European Journal of Foreign Language Teaching, 3*(3), 33-51. https://doi.org/10.5281/zenodo.1310086

Rood, M. M. S. J., & de Jong, N. H. (2023). A mixed-methods study on the effects of manipulated speaking anxiety on L2 utterance and L2 cognitive fluency. *Journal of the European Second Language Association, 7*(1), 60-74. https://doi.org/10.22599/jesla.96

Sandlund, E., Sundqvist, P., & Nyroos, L. (2016). Testing L2 talk: A review of empirical studies on second language oral proficiency testing. *Language and Linguistics Compass, 10*(1), 14-29. https://doi.org/10.1111/lnc3.12174

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

Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems,* 1-15. https://doi.org/10.1111/exsy.12667

Shirvan, M. E., & Talebzadeh, N. (2018). Foreign language anxiety and enjoyment in an imagined community. *Eurasian Journal of Applied Linguistics, 4*(2), 109-133. https://doi.org/10.32601/ejal.464043

Shohamy, E., Reves, T., & Bejarano, Y. (1986). Introducing a new comprehensive test of oral proficiency. *ELT Journal, 40*(3), 212-220. https://doi.org/10.1093/elt/40.3.212

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. https://doi.org/10.1037/0033-2909.86.2.420

Sipos, M. L., Kim, P. Y., Thomas, S. J., Adler, A. B., & Thomas, J. L**.** (2019). U.S. soldiers and foreign language school: Stressors and health. *Military Psychology, 31*(3), 238-248. https://doi.org/10.1080/08995605.2019.1582807

Skehan, P. (2013). Tasks and language performance assessment. In M. Bygate, M. Swain, & P. Skehan (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 191-210). Routledge.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93-120. https://doi.org/10.1111/1467-9922.00071

Song, B., & Lee, J. H. (2015). The effects of task practice and task type on L2 English oral performance. *Korean Journal of Applied Linguistics, 31*(2), 57-82. https://doi.org/10.1075/tblt.11.06hu

Susoy, Z. (2018). Individual vs. paired formats of oral proficiency exam: Does it matter in terms of scores and preference? *European Journal of Foreign Language Teaching*. *3*(1), 31-49. https://doi.org/10.5281/zenodo.1165470

Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics, 26*(3), 376-401. https://doi.org/10.1093/applin/ami013

Tavakoli, P., Nakatsuhara, F., & Hunter, A-M. (2017). Scoring validity of the Aptis speaking test: Investigating fluency across tasks and levels of proficiency. *Arags Research Reports Online,* British Council. http://www.britishcouncil.org/exam/aptis/research/publications/tavakoli

Teng, H.-C. (2007). *A study of task types for L2 speaking assessment*. [Paper Presentation]. Annual Meeting of the International Society for Language Studies (ISLS), Honolulu.

Toding, R. W., Yahrif, M., & Sirajuddin, S. (2021). Using unscripted role-play teaching activity to improve students' speaking ability. *PROJECT (Professional Journal of English Education)*, *4*(6), 1057-1066. https://doi.org/10.22460/project.v4i6.p1057-1066

Torky, S. (2006). *The effectiveness of a task-based instruction program in developing the English language speaking skills of secondary stage students* [Unpublished doctoral dissertation]. Ain Shams University.

Tsang, A. (2022). The relationships between EFL learners' anxiety in oral presentations, self-perceived pronunciation, and speaking proficiency. *Language Teaching Research, 29*(4), 1639-1659. https://doi.org/10.1177/13621688221102522

Vo, Y. (2018). *Effects of integrated and independent speaking tasks on learners' interactional performance*. [Unpublished master's thesis]. Minnesota State University.

Wigglesworth, G., & Taylor, L. (2009). Are two heads better than one? Pair work in L2 assessment contexts. *Language Testing, 26*(3), 325-339. https://doi.org/10.1177/0265532209104665

Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, *103*(4), 763-781. https://doi.org/10.1111/modl.12590

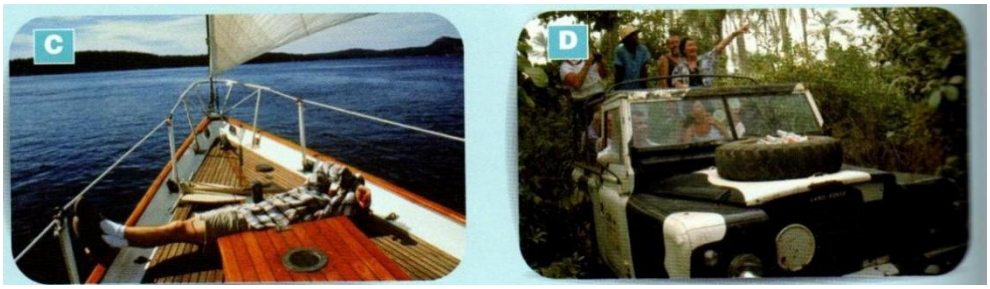
Zhang, W., & Liu, M. (2013). Evaluating the impact of oral test anxiety and speaking strategy use on oral English performance. *Journal of Asia TEFL, 10*(2), 115-148.

Appendix A. Picture description task samples

Instructions: You will be shown three different sets of pictures, one at a time. For each set, please describe and compare the pictures in as much detail as possible. Try to speak continuously for about one minute.

Prompt 1

Prompt 2



Prompt 3

Appendix B. Role-play task sample

Scenario 1

Instructions: You and your partner will be given three different situations one by one. In each situation, you will have a specific role. Discuss the topic with your partner and try to reach a conclusion together. You will have about two minutes for each scenario.

Card A (candidate A)	**Card B (candidate B)**
Context: You're a customer at a clothes' shop. You want to buy a pair of black jeans. Task: -Ask for help. -Explain your request. -Ask for the color and the size you'd like to try on. -Ask for the available color. -Ask for the price. -Explain why you don't like the jeans you were given (color, size, price). -Thank and say goodbye.	Context: You are a shop assistant in a clothes shop. Respond to what the customer asks. You don't have the jeans in the colour the customer wants. Try to sell him a different colour. Task: -Respond the help request positively. -Explain that you have the right size but not the color the customer asks for. -Offer a different color on the same size. -Give the price (175 TL) after the customer asks. -Ask if the customer needs any further help. -Respond and say goodbye.

Appendix C. Rubrics

Overall scores assigned to the picture-cued description task

Linguistic competence subskills	Grammatical competence			Discourse competence	Pragmatic competence	Fluency	Task Achievement	Total score
	Grammar	Pronunciation	Vocabulary	Coherence				
Giving descriptions and making comparison	15(5x3)	15 (5x3)	15 (5x3)	15 (5x3)	15 (5x3)	15 (5x3)	10 (5x2)	100

Overall scores assigned to the role play task

Linguistic competence subskills	Grammatical competence			Discourse competence		Pragmatic competence	Fluency	Task Achievement	Total score
	Grammar	Pronunciation	Vocabulary	Coherence	Interactive communication				
Making /declining requests and interacting in social situations (greeting and leave taking)	10 (5x2)	15 (5x3)	10 (5x2)	10 (5x2)	15 (5x3)	15 (5x3)	15 (5x3)	10 (5x2)	100