Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 15 (4). 2025. 831-856. Published online: 04.08.2025
https://doi.org/10.14746/ssllt.46191
http://pressto.amu.edu.pl/index.php/ssllt

The influence of task type and learner language background on writing production and anxiety: A Bayesian linear mixed-effects analysis

Mahmoud Abdi Tabari ✉
University of Nevada, Reno, USA
https://orcid.org/0000-0002-8022-5415
mabditabari@unr.edu

Xinya Liang
University of Arkansas, Fayetteville, USA
https://orcid.org/0000-0002-2453-2162
xl014@uark.edu

Ágnes Albert
Eötvös Loránd University, Budapest, Hungary
https://orcid.org/0000-0001-8339-7119
albert.agnes@btk.elte.hu

Abstract

Despite growing interest in task-based language teaching (TBLT), limited empirical work has examined how different rhetorical task types influence second language (L2) writing development, especially in relation to affective variables, such as writing anxiety. Existing research in TBLT has largely focused on cognitive dimensions, often neglecting individual differences in learners' emotional responses. Moreover, Long's (2014) call to use first language (L1) data as a benchmark in TBLT remains underexplored, complicating the interpretation of L2 performance patterns. To address these gaps, we examined the impact of task type and writing anxiety on the written performance of 140 university students (70 L1 English speakers, 70 L2 English learners). Participants completed a writing anxiety questionnaire and performed four rhetorical tasks over four weeks. Essays were assessed using syntactic complexity, accuracy, lexical complexity, and fluency (CALF) indices. Bayesian linear mixed-effects modeling was used to analyze both linguistic and affective variables. Results revealed significantly higher levels of somatic anxiety, cognitive anxiety, and anxiety avoidance among L2 writers. In both groups, increased task complexity was associated with longer clauses and more sophisticated syntax, with L1 writers showing steady gains in accuracy and fluency. L2 writers, however, exhibited greater variability: accuracy spiked during the argumentative task, while lexical diversity declined. The findings highlight the interaction between cognitive task demands and affective factors in shaping L2 writing outcomes, offering insights for TBLT researchers and pedagogical practices in L2 writing.

Keywords: task type; learner group; anxiety; CALF; TBLT

1. Introduction

Anxiety is one of the most widely studied affective variables in second and foreign language learning, with early research dating back to the 1970s (e.g., Chastain, 1975; Kleinmann, 1977). The field has since evolved through several phases (MacIntyre, 2017), with recent studies increasingly focusing on specific instructional contexts, such as English as a second language (ESL) and English as a foreign language (EFL). Extensive evidence links language anxiety to overall achievement (Cakici, 2016; Thompson & Lee, 2014), and growing research has explored its impact on writing (Abdi Tabari & Goetze, 2024; Li et al., 2023). However, little is known about how anxiety affects performance on specific writing tasks, particularly within task-based language teaching (TBLT, Lambert et al., 2023). This gap has pedagogical implications. Writing instructors working within TBLT may benefit from understanding how task-specific anxiety shapes learners' performance. Although applying such insights can be complex, they may support the development of strategies to recognize and mitigate anxiety effects.

In the psycholinguistic approach to TBLT (Ellis, 2000), tasks are viewed as vehicles for specific cognitive processes. Robinson's (2003, 2005, 2015) triadic componential framework has been central to this perspective, highlighting how task complexity affects language production and learning. While most research has emphasized cognitive dimensions, affective variables, such as anxiety, may also play a critical role. Drawing on Lambert et al. (2023), we hypothesized that anxiety would exert a stronger influence on more cognitively demanding tasks. As task complexity increases, learners must allocate more mental resources; anxiety can deplete this limited capacity, hindering focus and processing (Vytal et al., 2013). Despite the relevance of Robinson's framework, few studies have examined how anxiety interacts with task complexity in writing (Johnson, 2017), and even fewer have compared native and non-native speakers on the same tasks (Abdi Tabari & Lee, 2024; Abdi Tabari et al., 2024; Foster & Tavakoli, 2009).

To address these gaps, this study investigated the writing performance of speakers of English as a first language (L1) and learners of English as a second language (L2) across four tasks of increasing cognitive complexity: narrative, expository, expo-argumentative, and argumentative. Participants also completed a writing anxiety questionnaire. L1 writers in our study provided a baseline for interpreting L2 performance. Moreover, while writing anxiety may affect both groups, it is likely heightened in L2 learners due to differences in proficiency, self-perceptions, and beliefs about language use (Cheng, 2002). We examined how task complexity, anxiety, and learner group influenced four core dimensions of writing: syntactic complexity, accuracy, lexical complexity, and fluency (CALF).

2. Literature review

2.1. The role of tasks in written production

TBLT literature has extensively examined challenges related to task design, sequencing, and implementation, focusing on two key hypotheses: the limited attentional capacity (LAC) model (Skehan, 2001, 2014) and the cognition hypothesis (CH; Robinson, 2003, 2005). Both models explore the relationship between task features and CALF in language production, highlighting the role of cognitive processes and attentional resources. Skehan's LAC model suggests that L2 learners prioritize specific CALF aspects during demanding tasks, while Robinson's CH posits that task cognitive complexity influences L2 performance, potentially fostering skill development and redirecting attention to accuracy and complexity. Robinson (2015) advocates sequencing tasks based on cognitive complexity. Both scholars list task features affecting L2 writing performance, including content availability (Révész et al., 2017), topic familiarity (Li & Yu, 2023), pre-task planning (Abdi Tabari, 2022), task characteristics (Levkina & Gilabert, 2012), and reasoning demands (Rahimi, 2018). These studies emphasize the role of task types and other features in shaping L2 writing performance.

Task types elicit distinct linguistic features and introduce varying information-processing pressures (Barber, 2015), yet TBLT literature presents inconsistent findings on their effects on written production (Kim & Pae, 2021). This inconsistency is particularly pronounced with the rhetorical task (RT), recognized in

L1 and L2 writing as both rhetorical and discourse mode (Weigle, 2002). RT includes narration, description, exposition, and argumentation, each with distinct cognitive demands. The complexity progresses from narration to argumentation, with argumentation requiring the most reasoning (Yang, 2014).

Studies consistently show that both L1 and L2 writers produce more complex texts in argumentative tasks than in narratives (Beers & Nagy, 2009; Bi, 2020; Lu, 2011). These findings align with Robinson's (2003, 2005) CH, which asserts that more cognitively complex tasks lead to more complex language output. However, no TBLT study has compared L1 and L2 writers' performance on RTs, and research on RTs has been limited to just a few writing tasks. Therefore, addressing these gaps in task-based L2 writing research is crucial.

2.2. Anxiety and writing

In addition to task design and implementation factors, learners' language performance is shaped by cognitive, affective, and conative traits known as individual difference (ID) variables. One affective ID variable, anxiety, has been studied in second language acquisition (SLA) for nearly 50 years (Chastain, 1975; Kleinmann, 1977). Initially conceptualized as a trait, anxiety was later seen as a situation-specific emotion in SLA (MacIntyre & Gardner, 1994). Labeled as *foreign language learning anxiety*, it has been linked to difficulties and a lack of success in second/foreign language learning (Horwitz et al., 1986). Defined as "the feeling of tension and apprehension specifically associated with second language contexts" (MacIntyre & Gardner, 1994, p. 284), foreign language learning anxiety negatively affects input, processing, and output phases of learning (MacIntyre & Gardner, 1994). While the *Foreign Language Classroom Anxiety Scale* (FLCAS, Horwitz et al., 1986) mainly focuses on speaking (Gkonou, 2011), measures for other skills have been developed, such as Cheng's (2004) *Second Language Writing Anxiety Inventory* (SLWAI), used in this study.

Although anxiety is an ID variable affecting L2 learning and production, writing itself can provoke anxiety in both L1 and L2 contexts. Due to its cognitively demanding nature and high working memory demands (Kim, 2020), writing often causes anxiety, even in the first language. Daly and Miller's (1975) first writing anxiety scale focused on writing apprehension in L1 contexts, coinciding with early research on anxiety in second/foreign language learning (Chastain, 1975; Kleinmann, 1977). Their scale, later adapted for L2 contexts (e.g., Cheng et al., 1999), viewed writing apprehension as a one-dimensional construct, primarily about disliking and avoiding writing. In contrast, Cheng's (2004) SLWAI adopts a multidimensional view, differentiating between somatic anxiety (physiological symptoms), cognitive anxiety (worry), and avoidance behavior (procrastination), thus offering a more comprehensive view of the manifestation of anxiety during writing. Given that both L1 and L2 writing are complex cognitive tasks, this approach allows for a more accurate understanding of anxiety effects across languages.

We chose Cheng's (2004) SLWAI for our study because it measures general anxiety levels and isolates specific sources of anxiety, providing a nuanced view of how anxiety affects writing in both L1 and L2 contexts. The SLWAI has been widely validated and used in research on second language writing anxiety (Cheng, 2004; Cheng et al., 1999), making it a reliable tool for assessing anxiety in L2 writing. Cheng's model accounts for the complex interplay between anxiety and performance as most studies show a negative correlation between the two (Abdi Tabari et al., 2025; Busse et al., 2023; Cheng, 2002; Cheng et al., 1999). Using the SLWAI enables us to assess how anxiety influences writing in both L1 and L2 contexts, offering insights into how anxiety is uniquely experienced and how it affects writing performance. Thus, the SLWAI provides a detailed, multi-dimensional understanding of writing anxiety, central to exploring its impact on L1 and L2 writing performance.

2.3. Anxiety and performance on written tasks

Despite numerous studies on anxiety and L2 writing, research examining the relationship between learners' performance on specific task types and anxiety – particularly studies using CALF measures – remains limited. A few studies have compared anxiety levels across spoken and written tasks, yielding mixed results. Baralt and Gurzynski-Weiss (2011) found no difference in anxiety levels between face-to-face oral and written synchronous computer-mediated communication (SCMC) tasks among intermediate-level Spanish learners. In contrast, Côté and Gaffney (2021) reported that beginner French learners experienced less anxiety during a written task, leading to more turns and words produced in SCMC compared to face-to-face (FTF) tasks. Trebits (2016) explored different task modalities and cognitive complexity in her study with Hungarian high school students. She found that output anxiety was linked to lower accuracy in cognitively more complex spoken tasks, supporting Robinson's (2003, 2005) claim that individual differences (IDs) have a greater impact on performance in more complex tasks. Additionally, Trebits (2016) found that processing anxiety (related to mental effort during task execution) positively correlated with vocabulary and syntactic complexity on cognitively less complex spoken tasks, while in writing, anxiety negatively impacted accuracy, vocabulary, and syntactic complexity. These effects were more pronounced in the cognitively more complex task.

We identified one empirical study using Chen's (2004) writing anxiety scale to investigate the relationship between somatic anxiety, cognitive anxiety, avoidance behavior and learners' performance on writing tasks. Zabihi et al. (2020) designed two tasks – narration (less complex) and argumentation (more complex) – and compared the effects of task complexity on writing performance. Anxiety was evident across both tasks, though its effects differed by task complexity. On the less complex narrative task, cognitive anxiety negatively impacted accuracy, and somatic anxiety affected one accuracy measure. On the more complex argumentative task, anxiety had a stronger negative impact, with cognitive anxiety correlating significantly with complexity, accuracy, and fluency. Avoidance behavior was also negatively associated with fluency and complexity measures, similarly to Trebits (2016), where anxiety effects were more pronounced on cognitively complex tasks.

These findings suggest that while anxiety is generally viewed as an ID variable with a detrimental effect on language performance, there is a lack of research on the relationship between writing anxiety and specific task types in L1 and L2 writing. Further investigation is needed on how somatic anxiety, cognitive anxiety, and avoidance behavior relate to L1 and L2 learners' writing performance. Additionally, Robinson's (2015) assertion that learner factors more strongly influence performance on cognitively complex tasks should be tested empirically.

3. The current study

Building on previous discussions and addressing identified gaps, the current study aimed to achieve two main objectives. First, we sought to identify differences in writing anxiety between L1 and L2 learners across subscales – somatic anxiety, avoidance behavior, and cognitive anxiety – and explain these differences. L2 learners were expected to experience higher anxiety due to lower proficiency and negative self-perceptions despite writing being cognitively challenging also for L1 learners. Second, we investigated the interaction between task types and learner groups on CALF measures, considering writing anxiety as a controlled variable. Based on Robinson's (2003) CH, we hypothesized that IDs would have a greater impact on more cognitively complex tasks, with proficiency differences moderating performance. The three RTs represented varying levels of cognitive complexity. To address these objectives, the study formulated the following research questions:

1. Is there a discernible difference between L1 and L2 learner groups regarding various subscales of writing anxiety, namely somatic anxiety, avoidance behavior, and cognitive anxiety? If so, how do they differ? (RQ1)
2. What is the interaction effect between task types and learner groups on CALF measures, considering writing anxiety as a controlled variable? (RQ2)

4. Methodology

4.1. The learning context and participants

This study recruited 140 learners (70 L1 and 70 L2) from a large university in the US using convenience sampling, with participants enrolled in six sections of a first-year writing course. These sections included three for mainstream learners and three for international/multilingual (L2) learners. The course aimed to help students adapt their writing to various genres and audiences, focusing on rhetorical processes such as audience, purpose, and occasion. The instruction included strategies for planning, drafting, and revising. All participants attended weekly 2.5-hour sessions, following a standardized curriculum approved by the Core Writing program. The course was taught by two instructors using the TBLT approach.

Participants, aged 19 to 21, included 93 females and 47 males. The 70 international students came from diverse linguistic backgrounds, such as Chinese, Korean, Japanese, Russian, French, Spanish, Portuguese, and Persian. Like their L1 peers, they had completed 2-5 academic writing courses before the study and were identified as having C1 writing proficiency according to the *Common European Framework of Reference for Languages* (CEFR, 2001, 2020). Ethical consent forms were signed by all participants, with approval from the Institutional Review Board (IRB) of the university.

4.2. Instruments

4.2.1. Writing tasks

The study included four rhetorical writing tasks: narration, exposition, expo-argumentation, and argumentation. These tasks were chosen for their relevance to first-year writing courses and alignment with the Core Writing Program curriculum. Each task was connected to the course theme, “writing and technology,” and designed to increase cognitive challenges, progressing from narration to argumentation. This progression mirrors real-world writing demands, promoting adaptability and versatility. The narrative task focused on personal storytelling, while the expository task was informational and non-argumentative. The expo-argumentative task combined exposition and argumentation, and the argumentative task emphasized persuasion with evidence. Each task placed different cognitive and rhetorical demands on participants.

4.2.2. Writing anxiety questionnaire

The study adapted Cheng's (2004) SLWAI, incorporating its three anxiety subscales: somatic anxiety, cognitive anxiety, and avoidance behavior. The original SLWAI comprises 22 items, with scores ranging from 22 to 132, where higher scores indicate elevated anxiety. A key modification replaced the original 5-point Likert scale with a 6-point scale, where respondents expressed agreement from "strongly disagree" (1) to "strongly agree" (6). This change was based on research suggesting that 6-point scales provide better discrimination and reliability, particularly for studies with multiple variables (Chomeya, 2010), and they often show more favorable normality than 5-point scales. Further modifications ensured the instrument was more applicable to mainstream participants, including revisions to item wording to make them more relevant to the broader context while maintaining the integrity of the subscales. The internal consistency of the modified SLWAI was high (α = .95 for somatic anxiety, α = .98 for avoidance behavior, and α = .92 for cognitive anxiety).

Examples of items for each subscale were as follows. For somatic anxiety, items included "I feel my heart pounding when I write English compositions under time constraints" and "I tremble or perspire when I write English compositions under time pressure." For cognitive anxiety, examples were "My mind often goes blank when I start to work on an English composition," and "I worry about making mistakes when writing in English." For avoidance behavior, items included "I usually avoid writing English compositions" and "I would excuse myself if asked to write English compositions." These examples illustrate how the subscales measured distinct aspects of language-related anxiety, including physical reactions (somatic anxiety), cognitive concerns (cognitive anxiety), and avoidance tendencies (avoidance behavior).

4.3. Data collection procedures

First, we contacted course instructors to explain the objectives and procedures of the study, seeking their collaboration and inquiring about their willingness to integrate the four writing task types into their classes. Once confirmed, we scheduled visits from Week 4 to Week 8 of the term. During the first visit (Week 4), participants were briefed on the study goals, their voluntary participation rights, and the confidentiality of their information. All participants consented by signing the forms, with no withdrawals. In the same session, international students completed the IELTS writing task 2, serving as a pretest to assess their writing proficiency levels. The following week, participants completed the writing anxiety questionnaire, followed by the written narrative task using a word processor. Over the next three weeks, participants completed expository, expo-argumentative, and argumentative tasks, each separated by a one-week interval in alignment with their class schedule. In consultation with instructors, we set time limits for each task: narrative (35 minutes), expository (40 minutes), expo-argumentative (45 minutes), and argumentative (50 minutes). After task completion, participants' essays were analyzed using CALF measures. Finally, the questionnaire data were coded and analyzed comprehensively.

4.4. CALF measurement

To address the complex nature of CALF, we opted for non-repeated measures to mitigate potential multicollinearity. Our choice of measures followed two principles: first, we used measures from prior TBLT studies to enhance the generalizability of our findings and support L2 writing TBLT research; second, we selected distinct measures to capture various dimensions within each construct.

We utilized two natural language processing *(NLP) tools – L2 Syntactic Complexity* Analyzer (Lu, 2010) and *Coh-Metrix* (McNamara et al., 2014) – to select five out of 14 complexity measures. After screening for collinearity, these measures were categorized based on Lu's (2011) framework: *mean length of sentence (MLS)* and *mean length of clause* (MLC) for "length of the production unit," representing clause-level and phrase-level complexity, respectively. Additionally, we calculated *T-units per sentence* (T/S) for coordination, *clauses per T-unit* (*C/T*) for subordination, and *complex nominals per clause* (CN/C) for phrasal sophistication. In addition, accuracy was assessed using two indices: first, the proportion of *error-free clauses* (EFCs), which included lexical, morphological, and syntactic errors, excluding spelling, capitalization, and punctuation errors; second, the proportion of *correct verb forms* (CVFs), which measured verb usage accuracy, including subject-verb agreement, modal verbs, tense, and aspect. *Textual lexical diversity* (MTLD) and *the log frequency of content words* (LCW) were used to assess lexical complexity. MTLD measures "the mean length of sequential word strings in a text that maintains a given TTR value" (Mazgutova & Kormos, 2015, p. 5). LCW, a measure of lexical sophistication, was calculated as the average log frequency of content words in the CELEX database (Graesser et al., 2004). LCW was preferred over frequency band measures for its suitability in tracking small and large gains in individual writings and its reliability in assessing lexical sophistication (Kormos, 2011). Finally, *words per minute* (*W/M*) was used as a fluency measure, following Kellogg's (1996) recommendation as "an average measure of fluency" (p. 65) and recognized by Abdi Tabari (2023) as a reliable indicator of L2 writing fluency.

4.5. Data analysis

The analysis examined three anxiety subscales and ten CALF outcomes using *R* software (R Core Team, 2022). Bayesian linear mixed-effects models were employed with the "brms" package (Bürkner, 2017), which uses the Stan program for Bayesian estimation (Stan Development Team, 2023). For each anxiety subscale (somatic anxiety, avoidance behavior, and cognitive anxiety), the subscale scores were obtained by averaging the items. The learner group (L1 and L2) was included as the independent variable. The learner difference was analyzed as fixed effects, with the intercept and group modeled as random effects. For each of the ten CALF measures (MLS, MLC, T/S, C/T, CN/C, EFCs, CVFs, MTLD, LCW, and WM) as the outcome variable, independent variables included written task (narrative, expository, expo-argumentative, argumentative), learner group (L1 and L2), the interaction between task and group, and the three anxiety subscales. This model aimed to reveal the main and interaction effects between task and group on CALF measures while controlling for anxiety levels. Interactions between task type and anxiety measures were tested but were mostly non-significant and excluded from the final model. The intercept and tasks were included as random effects as tasks were measured within the same student.

Bayesian linear mixed models were chosen due to the small sample sizes and violations of normality assumptions for some outcome variables. For example, cognitive anxiety showed a negatively skewed distribution, which the posterior distribution successfully approximates. Bayesian estimation offers benefits such as incorporating prior knowledge, not relying on distributional assumptions, and accommodating small sample sizes (Lee, 2007). In Bayesian estimation via Stan, posterior distributions are approximated using the Hamiltonian Monte Carlo method (Neal, 2011) and its extension, the *No-U-Turn Sampler* (NUTS) (Hoffman & Gelman, 2014). Four chains were employed, each with 1,000 warm-up draws, 2,000 sampling iterations, and 1,000 post-warmup draws. Convergence was assessed using Rhat, the potential scale reduction factor on split chains, with Rhat < 1.1 indicating convergence (Gelman & Rubin, 1992). Default noninformative priors were applied for Bayesian estimation due to the limited research in this area. Specifically, flat or uniform priors were assigned for fixed effects (intercept, slope), LKJ priors for correlations, and student-t priors for standard deviations.

To interpret the effect of an independent variable on the outcome, model parameters were summarized using the mean and standard deviation of the posterior distribution along with two-sided credible intervals (denoted "l-95% CI" for the lower bound and "u-95% CI" for the upper bound). Credible intervals serve to determine significance where, if the interval captures zero, it indicates a non-significant effect, and vice versa. In addition, two R^2 effect size measures were provided to indicate the proportion of variance in the outcome explained by the independent variables. The first is the R_1^2, which is conditional on the random effects, and the second is the R_2^2, setting all random effects to zero and conditioned on an "average" random effects level. These measures are akin to the R^2 proposed for the linear mixed models estimated using frequentist methods (Nakagawa & Schielzeth, 2013).

5. Results

5.1. Anxiety

To provide a clearer understanding of the data distribution, the results for the different types of anxiety are described below. Somatic anxiety (M = 3.43, SD = 1.74) displayed a bimodal distribution, with most responses clustering around categories 2 and 5. Avoidance behaviors (M = 3.38, SD = 0.38) were approximately normally distributed. Cognitive anxiety (M = 3.35, SD = 0.37) demonstrated a negative skew. Although not every outcome exhibited a normal distribution, the posterior distributions effectively represented the data distribution for each outcome, indicating a good model-data fit.

Table 1 Fixed-effects estimates for anxiety

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Somatic							
Intercept	1.72	0.04	1.65	1.8	1.01	924	1077
L2	3.41*	0.06	3.3	3.52	1	1274	2651
Avoidance							
Intercept	3.14	0.04	3.07	3.22	1.02	789	775
L2	0.48*	0.05	0.38	0.58	1.01	1232	1806
Cognitive							
Intercept	3.22	0.05	3.12	3.32	1.01	715	1113
L2	0.25*	0.06	0.13	0.36	1.01	776	1331

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 1 shows the learner group difference in somatic anxiety, avoidance behaviors, and cognitive anxiety measures. Compared to the L1 group, the L2 group showed a significantly higher level of somatic anxiety β = 3.41, 95% CI = [3.30, 3.52]), more avoidance behavior (β = 0.48, 95% CI = [0.398, 0.58]), and a higher level of cognitive anxiety (β = 0.25, 95% CI = [0.13, 0.36]). The R_1^2 and R_2^2 were .737 and .723 respectively, representing the proportion of variances in somatic anxiety explained by the learner group difference when considering random effects versus not considering them. Similarly, the R_1^2 and R_2^2 were .565 and .306 for avoidance behavior, and .601

and .102 for cognitive anxiety. Thus, somatic anxiety distinguished between learner groups better than avoidance behaviors or cognitive anxiety.

5.2. Syntactic complexity

Among the five measures of syntactic complexity, MLS (*M* = 15.82, *SD* = 2.9), MLC (*M* = 8.81, *SD* = 1.32), T/S (*M* = 1.31, *SD* = 0.22), and C/T (*M* = 1.56, *SD* = 0.16) were approximately normal, while CN/C (*M* = 1.35, *SD* = 0.23) was slightly positively skewed. These distributions were all well represented by the posteriors. The correlation among these measures was low, ranging from 0.13 to 0.31. Regarding MLS, Table 2 shows the influence of task type and learner group on the MLS measure, accounting for anxiety levels. The interactions between task type and learner groups were found to be non-significant, leading us to examine the main effects. Utilizing the argumentative task as the reference group, both the expository and narrative tasks significantly decreased the MLS, with the narrative task displaying the lowest MLS. The effects of the expo-argumentative task, learner group, and anxiety level on MLS were not found to be significant. The R_1^2 and R_2^2, at .353 and .268, respectively, denote the percent of variances in MLS explained by the model when considering versus not considering random effects.

Table 2 Fixed-effects estimates for MLS

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	17.69	1.35	15.12	20.41	1	4002	3441
Expo-argumentative	-0.43	0.38	-1.17	0.31	1	2344	2592
Expository	-1.46*	0.39	-2.25	-0.69	1	1851	2622
Narrative	-2.38*	0.39	-3.14	-1.63	1	1700	2304
L2	-2.01	1.26	-4.4	0.58	1	951	711
Somatic	-0.51	0.36	-1.24	0.17	1	953	603
Avoidance	0.14	0.5	-0.85	1.14	1	2812	2993
Cognitive	0.34	0.43	-0.51	1.17	1	2148	2494
Expo-argumentative × L2	0.76	0.54	-0.3	1.83	1	2278	2676
Expository × L2	0.86	0.56	-0.22	1.96	1	2010	2675
Narrative × L2	0.88	0.55	-0.18	1.95	1	1602	2437

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 3 depicts the impact of task type and learner group on the MLC measure, considering anxiety levels. The only significant interaction was between task type (argumentative vs. narrative) and group (L1 vs L2). The L1/L2 learner difference in MLC for the argumentative task was about 1.5, compared to that for the narrative task, about 0.4. The three anxiety measures did not have a significant influence on MLC. The variances explained were R_1^2 = .485 and R_2^2 = .219.

Table 3 Fixed-effects estimates for MLC

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	9.99	0.6	8.79	11.15	1	2613	2896
Expo-argumentative	-0.33	0.21	-0.73	0.07	1	873	1511
Expository	-0.82*	0.21	-1.24	-0.4	1.01	818	861
Narrative	-1.66*	0.18	-2	-1.3	1	1134	1742
L2	-1.54*	0.56	-2.63	-0.42	1	1480	2376
Somatic	0.02	0.15	-0.29	0.31	1	1740	2439
Avoidance	-0.06	0.23	-0.49	0.38	1	1881	1804
Cognitive	0.06	0.19	-0.31	0.42	1	2037	2066
Expo-argumentative × L2	0.14	0.29	-0.42	0.7	1	1179	2040
Expository × L2	0.49	0.3	-0.08	1.08	1	1151	2030
Narrative × L2	1.18*	0.25	0.67	1.66	1	1580	1973

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 4 illustrates the impact of task type and learner group on the T/S measure, considering anxiety levels. The interactions between task type and learner groups were not statistically significant. Using the argumentative task as the reference, both the expository and narrative tasks notably reduced the T/S, with the narrative task showing the lowest T/S. L2 learners exhibited a statistically significant .34 lower T/S than L1 learners. The effects of the expo-argumentative task and anxiety level on T/S were not found to be statistically significant. The R_1^2 and R_2^2 were .201 and .101, respectively.

Table 4 Fixed-effects estimates for TS

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.22	0.1	1.03	1.41	1	2779	2886
Expo-argumentative	-0.06	0.04	-0.13	0.01	1	1510	1857
Expository	-0.13*	0.03	-0.2	-0.07	1	1712	2452
Narrative	-0.17*	0.03	-0.23	-0.1	1	1669	2897
L2	-0.34*	0.09	-0.52	-0.16	1	1931	2311
Somatic	0.04	0.03	-0.01	0.09	1	2107	2059
Avoidance	0.04	0.04	-0.03	0.11	1	2799	2999
Cognitive	0.02	0.03	-0.04	0.08	1	2483	2412
Expo-argumentative × L2	0.02	0.05	-0.08	0.12	1	1339	2022
Expository × L2	0.01	0.05	-0.08	0.09	1	1489	2302
Narrative × L2	-0.01	0.05	-0.1	0.08	1	1480	2645

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 5 shows the influence of task type and learner group on the C/T measure, accounting for anxiety levels. The interactions between task type and learner groups were non-significant. Examining the main effects, compared to the argumentative task, both the expository and narrative tasks significantly decreased the C/T, with the narrative task displaying the lowest C/T. The effects of the expo-argumentative task, learner group, and anxiety level on C/T were not found to be significant. The variances explained were $R_1^2 = .353$ and $R_2^2 = .268$.

Table 5 Fixed-effects estimates for CT

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.69	0.08	1.54	1.85	1	4086	3202
Expo-argumentative	-0.03	0.03	-0.08	0.03	1	1439	705
Expository	-0.05*	0.02	-0.1	-0.01	1	2102	2417
Narrative	-0.12*	0.02	-0.16	-0.07	1	2250	2510
L2	-0.03	0.07	-0.17	0.11	1	2390	2548
Somatic	-0.01	0.02	-0.05	0.03	1	2540	2742
Avoidance	0	0.03	-0.06	0.06	1	2981	2701
Cognitive	-0.01	0.03	-0.06	0.04	1	3172	2975
Expo-argumentative × L2	-0.01	0.04	-0.08	0.07	1	1728	1467
Expository × L2	-0.05	0.04	-0.12	0.02	1	1705	2162
Narrative × L2	0.04	0.04	-0.03	0.11	1	2113	2387

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero).

Table 6 presents the impact of task type and learner group on the CN/C measure, considering anxiety levels. The interactions between task type and learner groups were not statistically significant. Using the argumentative task as the reference group, only the narrative task showed a significantly lower CN/C of .18. All other effects were not significant. The R_1^2 and R_2^2 were .380 and .179, respectively.

Table 6 Fixed-effects estimates for CNC

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.57	0.11	1.35	1.79	1	4168	3238
Expo-argumentative	-0.07	0.04	-0.14	0	1	3104	3015
Expository	-0.01	0.03	-0.07	0.05	1	2989	3198
Narrative	-0.11*	0.03	-0.17	-0.05	1	3185	3011
L2	-0.18	0.1	-0.38	0.02	1	3023	2788
Somatic	-0.01	0.03	-0.06	0.05	1	2724	2858
Avoidance	0	0.04	-0.09	0.08	1	3294	1602
Cognitive	-0.02	0.04	-0.08	0.05	1	3142	2331
Expo-argumentative × L2	0.04	0.05	-0.06	0.14	1	2731	1800
Expository × L2	-0.06	0.05	-0.15	0.03	1	2958	2853
Narrative × L2	0.04	0.04	-0.04	0.13	1	3203	3142

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

5.3. Accuracy

Both EFCs ($M = 0.58$, $SD = 0.11$) and CVFs ($M = 0.59$, $SD = 0.11$) were normally distributed. They were weakly correlated with a correlation of 0.25. Table 7 illustrates the impact of task type and learner group on EFCs while controlling anxiety levels. The interaction between tasks (argumentative vs. narrative) and learner groups (L1 vs. L2) was found to be significant. The mean group difference in EFCs was relatively large at approximately .7 for the argumentative task, decreasing to about .2 for the narrative task. Anxiety levels did not exhibit an impact on EFCs. The variances in EFCs explained by the model were .517 (R_1^2, with random effects) and .112 (R_2^2, without random effects).

Table 7 Fixed-effects estimates for EFCs

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.56	0.05	0.45	0.66	1	2443	2774
Expo-argumentative	-0.02	0.02	-0.06	0.02	1	1094	1807
Expository	-0.05*	0.02	-0.09	-0.01	1	1124	1834
Narrative	-0.09*	0.02	-0.12	-0.06	1	1491	2436
L2	-0.07	0.05	-0.16	0.03	1	1661	2071
Somatic	0	0.01	-0.03	0.02	1	1802	2114
Avoidance	0.03	0.02	-0.01	0.07	1	1803	2595
Cognitive	0	0.02	-0.03	0.03	1	2166	2789
Expo-argumentative × L2	-0.02	0.03	-0.07	0.04	1	1378	2047
Expository × L2	-0.01	0.03	-0.06	0.05	1	1265	1666
Narrative × L2	0.05*	0.02	0.01	0.09	1	1488	2186

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 8 Fixed-effects estimates for CVFs

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.58	0.05	0.48	0.69	1	5210	3118
Expo-argumentative	-0.02	0.02	-0.05	0.01	1	2732	2489
Expository	-0.04	0.02	-0.07	0	1	2434	2986
Narrative	-0.07*	0.02	-0.1	-0.04	1	2983	2974
L2	-0.13*	0.05	-0.22	-0.04	1	3552	3043
Somatic	0.02	0.01	-0.01	0.04	1	3914	3114
Avoidance	0.01	0.02	-0.03	0.05	1	3383	3028
Cognitive	0.01	0.02	-0.03	0.04	1	4155	2635
Expo-argumentative × L2	0	0.02	-0.05	0.05	1	2633	2657
Expository × L2	0	0.02	-0.05	0.04	1	2282	2743
Narrative × L2	0.03	0.02	-0.01	0.07	1	2710	2862

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 8 displays the impact of task type and learner group on CVFs, considering anxiety levels. Non-significant interactions prompted an exploration of the main effects. Using the argumentative task as the reference, both expository and narrative tasks significantly decreased CVFs, with the narrative task having the lowest CVFs. Other effects were not significant. The variances explained were R_1^2 = .353 and R_2^2 = .268.

5.4. Lexical complexity

Both MTLD (*M* = 73.9, *SD* = 10.87) and LCW (*M* = 1.74, *SD* = 0.14) were normally distributed. They were weakly correlated with a correlation of 0.14. Table 9 depicts the impact of task type and learner group on the MTLD measure while accounting for anxiety. The three interactions between paired task types and learner groups were significant. Specifically, the most substantial difference between L1 and L2 groups was observed in completing the argumentative task (~17 point difference in MTLD), followed by the narrative task (~10 point difference), the expo-augmentative task (~7 point difference), and the expository task (~5 point difference). Anxiety levels did not contribute to explaining the variance in MTLD. The R_1^2 and R_2^2 were .296 and .110, respectively.

Table 9 Fixed-effects estimates for MTLD

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	80.86	5.8	69.46	92.21	1	2428	2597
Expo-argumentative	-2.07	1.83	-5.69	1.59	1	1028	1103
Expository	-5.67*	1.6	-8.82	-2.51	1	1374	1487
Narrative	-2.78	1.61	-5.93	0.41	1	1233	1327
L2	-16.35*	5.43	-27.03	-5.97	1	1356	1899
Somatic	1.25	1.55	-1.76	4.26	1	1448	2123
Avoidance	-1.15	2.15	-5.31	3.02	1	1442	1967
Cognitive	0.04	1.8	-3.55	3.55	1	1545	1976
Expo-argumentative × L2	8.74*	2.62	3.51	14	1	1153	1313
Expository × L2	11.25*	2.23	7	15.65	1	1412	2337
Narrative × L2	6.15*	2.23	1.62	10.6	1	1260	1983

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 10 Fixed-effects estimates for LCW

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.72	0.07	1.59	1.87	1	4155	3028
Expo-argumentative	-0.04	0.02	-0.08	0	1	3401	2951
Expository	-0.09*	0.02	-0.12	-0.05	1	3584	3097
Narrative	-0.07*	0.02	-0.12	-0.03	1	3066	3105
L2	-0.11	0.07	-0.25	0.01	1	3251	2992
Somatic	0.01	0.02	-0.02	0.05	1	3449	3136
Avoidance	-0.01	0.03	-0.06	0.04	1	3298	2839
Cognitive	0.03	0.02	-0.01	0.08	1	3412	2789
Expo-argumentative × L2	0.02	0.03	-0.03	0.08	1	3250	3166
Expository × L2	0.08*	0.03	0.03	0.13	1	3253	3356
Narrative × L2	0.04	0.03	-0.02	0.11	1	3323	3124

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

Table 10 presents the results for LCW. Only the interaction between tasks (argumentative vs. expository) and learner groups (L1 vs. L2) was significant, though the effect was small. Specifically, the learner difference in LCW for the argumentative task was about 0.11, while for the expository task it was about 0.04. The variances in LCW explained by the model were .302 (R_1^2, with random effects) and .060 (R_2^2, without random effects), respectively.

5.5. Fluency

Regarding the product-based fluency metric, the results revealed that WM (M = 10.33, SD = 1.35) displayed a normal distribution. Table 11 presents the effect of task type and learner group on the WM measure, adjusting for anxiety levels. To further explore these effects, the interactions between task type and learner group were significant. The most significant difference between L1 and L2 groups was observed in the argumentative task (~1.7 point difference in WM), followed by the expo-augmentative task (~1.1 point difference), the expository task (~0.8 point difference), and the narrative task (~0.7 point difference). Anxiety levels did not influence WM. The variances explained were R_1^2 = .315 and R_2^2 = .158.

Table 11 Fixed-effects estimates for WM

	Estimate	Est. error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	11.16	0.65	9.88	12.47	1	3888	2911
Expo-argumentative	-0.83*	0.21	-1.24	-0.41	1	1932	2650
Expository	-1.21*	0.2	-1.6	-0.82	1	1930	2490
Narrative	-1.27*	0.19	-1.64	-0.87	1	2313	2564
L2	-1.74*	0.61	-2.96	-0.56	1	2235	2477
Somatic	0.03	0.17	-0.31	0.37	1	2307	2345
Avoidance	0.22	0.25	-0.26	0.71	1	2242	2158
Cognitive	-0.09	0.21	-0.48	0.32	1	2396	2593
Expo-argumentative × L2	0.6*	0.29	0.03	1.17	1	1967	2680
Expository × L2	0.89*	0.29	0.32	1.44	1	2182	2828
Narrative × L2	1.01*	0.27	0.49	1.56	1	2280	2710

Note. Est. error = posterior standard deviation; l-95% CI = lower bound of 95% CI; u-95% CI = upper bound of 95% CI; Rhat = potential scale reduction factor; Bulk_ESS = estimated bulk effective sample size measuring sampling efficiency in Bayesian estimation; Tail_ESS = estimated tail effective sample size for 5% and 95% quantiles; * = significant effect (95% credible interval does not cover zero)

6. Discussion

RQ1 asked whether there was a discernible difference between L1 and L2 learner groups regarding various subscales of writing anxiety, namely somatic anxiety, avoidance behavior, and cognitive anxiety. If so, how do they differ? When examining the differences in writing anxiety between the L1 and L2 groups, it was discovered that L2 learners experienced more anxiety in connection with their writing tasks compared to their native-speaker peers. This finding may not be surprising as factors such as language proficiency and attitudes toward the L2 might have contributed to their anxiety, in addition to the negative affective state induced by the cognitively demanding nature of the writing tasks (Cheng, 2002). Among the three facets of writing anxiety (Cheng, 2004), the most significant difference between the two learner groups was observed in somatic anxiety, representing bodily symptoms and purely affective aspects. This suggests that L2 learners, through experiencing these physical symptoms, may have a stronger subjective feeling of anxiety than their L1 peers. Interestingly, the strongest correlations were identified between cognitive and behavioral anxiety scales and task performance measures in Zabihi et al.'s (2020) study, suggesting that these are linked to task performance more emphatically. Since the differences between the L1 and L2 learner groups were less pronounced on those scales in our study, we hypothesize that anxiety might have affected the performance of L1 and L2 learners in a similar manner, despite L2 learners subjectively feeling more anxious. If anxiety influences learners' performance on writing tasks similarly regardless of their native language background, it could explain the lack of significant interactions between the factors investigated in our study. As we are not aware of any previous research comparing anxiety levels between L1 and L2 speaker groups on the same set of writing tasks, our findings should be confirmed by future research, regardless of their seeming plausibility.

RQ2 concerned the interaction effect between task types and learner groups on CALF measures, considering writing anxiety as a controlled variable. Regarding syntactic complexity, the only measure where a significant interaction between task type and learner group could be found was the measure of MLC. Here, the gradual increase from the less complex to the most complex task was quite different in magnitude for the two groups. The L1 group seemed more adept at meeting task requirements, producing much longer clauses on the argumentative task than L2 learners. Although the group of L2 learners also increased their clause length, their relatively lower level of proficiency may have hindered them in meeting the requirements of the cognitively more complex tasks. This finding supports Robinson's (2003, 2005) claim that IDs, particularly language proficiency in this case, have a more pronounced effect on more complex tasks.

In the case of lexical complexity, both indices displayed interaction effects, indicating that participants' performance was simultaneously influenced by L1 and L2 group membership and task type. Interestingly, instead of a steady increase in the performance of the L1 group, there was a drop in lexical complexity in the expository task for both MTLD and LCW. It seemed as if the hypothesized cognitive complexity increase was not mirrored by accompanying lexical complexity gains in the case of this task. These results might mean that although expository tasks could be considered more challenging cognitively in general, solving this specific task did not require better lexis from the L1 group. The patterns were quite different in the case of the L2 group regarding lexical complexity. Specifically, the MTLD index was at its lowest in the argumentative task, while the LCW seemed relatively constant across the four tasks. This probably indicates L2 learners' inability to keep up with task requirements: For these L2 learners, increased cognitive complexity did not appear to provide enough impetus to push them toward creating lexically more complex texts. Although in the case of L1 speakers a simultaneous increase in syntax and lexis seemed to be the norm on cognitively more complex tasks, in language learners, the dissociation of syntax and lexis had been observed before, suggesting that at high levels of cognitive complexity, learners tend to focus on one of these aspects at the expense of the other (Skehan, 2014).

The general index of accuracy also showed interaction effects. In a similar manner to what has been observed in connection with syntactic complexity, L1 and L2 speakers behaved differently across the four tasks. In the case of the L1 group, a steady increase could be seen from the least complex narrative to the most complex argumentative task. Regarding the L2 group, the only notable increase in accuracy could be detected on the argumentative task. While the performance of the L1 group aligns with the predictions of the CH (Robinson, 2003, 2005, 2015), wherein the increase in cognitive complexity is met with parallel increases in accuracy, it appears that members of the L2 group might have struggled to meet the same task requirements. In their case, compared to the narrative task, only the argumentative task increased their accuracy. This finding also suggests that L2 learners might not always be able to meet the task demands posed by specific writing tasks, possibly due to limited attentional resources (Skehan, 2001, 2014).

Finally, the interaction between task type and group was also significant in the case of fluency, measured by WM. Here again, a steady increase was observed for the L1 group, while the gains in the L2 group were modest, resulting in a sizeable difference between the L1 and L2 groups on the cognitively most complex argumentative task. Therefore, it seems that cognitively more complex tasks not only enhanced L1 writers' performance qualitatively (via higher syntactic and lexical complexity and accuracy) but also boosted it quantitatively, encouraging them to write longer texts. Although the CH (Robinson, 2003, 2005) predicted reduced fluency for cognitively more complex tasks, this relationship was originally hypothesized to be true for spoken tasks. While it is plausible that the need to express more complex content probably affects speakers' ability to cope with real-time communication, thereby reducing fluency, time pressure appears to work somewhat differently in writing and probably causes less disruption in this regard. Also, more complex cognitive content might necessitate more explanation, resulting in longer texts. The tendency for parallel increases in cognitive complexity and output has already been identified in the case of spoken tasks (Albert, 2021). The increasing cognitive complexity of the RTs used in our study might have also urged both L1 and L2 writers to write more; however, L1 writers were more apt at meeting this task demand as well, resulting in an interaction effect. Since Swain's (2005) output hypothesis claims that generating output helps learners notice gaps in their knowledge, test hypotheses about language, and internalize linguistic forms, this finding might provide a further argument for using cognitively complex tasks in second/foreign language teaching.

7. Conclusion

This study examined the writing performance and anxiety of L1 and L2 English learners across tasks of increasing cognitive complexity. Although writing anxiety, measured using the SLWAI (Cheng, 2004), did not significantly interact with task type and was included as a covariate, notable differences emerged between L1 and L2 learners across the three subscales. The most pronounced difference between the groups was found in somatic anxiety, reflecting physical symptoms. However, prior research indicated that cognitive anxiety and avoidance behaviors strongly affect performance (Zabihi et al., 2020).

Regarding L2 task performance, findings generally support Robinson's (2003, 2005) CH as tasks with greater cognitive demands – from narrative to argumentative – were linked to improved performance in both groups, with some exceptions. For L1 learners, lexical complexity was unexpectedly lowest on the expository task. For L2 learners, lexical complexity did not improve with increased task complexity, and performance sometimes declined, suggesting difficulty in managing cognitive demands – a pattern also noted by Skehan (2014). Significant interactions between task type and learner group were found for syntactic and lexical complexity, accuracy, and fluency, indicating that L1 learners handled higher task demands more effectively, with the largest group differences observed in the argumentative task.

Although we expected anxiety to interact with task complexity and performance, results suggest that proficiency, particularly L1 versus L2 status, played a more decisive role. This underscores the impact of cognitive task complexity and proficiency on learners' ability to produce linguistically complex texts. The absence of significant interactions between anxiety and task complexity suggests that anxiety may affect performance independently of task demands. Future research could investigate this further using task-specific anxiety measures or qualitative data to better understand how learners experience and manage task-related anxiety.

This study has several limitations, the most notable being the absence of qualitative data, which could have provided a more nuanced understanding of learners' lived experiences with writing anxiety. Future research would benefit from integrating qualitative measures, such as interviews or open-ended surveys, to gain deeper insights into the complex ways in which anxiety affects writing behavior and performance. Additionally, administering a task-specific anxiety measure after each task could provide greater contextual specificity, allowing researchers to identify which aspects of writing tasks provoke anxiety. Such an approach would enable the development of more targeted interventions for students struggling with task-related anxiety. The lack of follow-up measures or in-depth interviews is another limitation as it leaves unexplored the coping strategies that students employ when facing writing anxiety and their perceptions of task difficulty. Understanding these factors is crucial, as it could reveal adaptive strategies that students use to manage anxiety, which are central to improving writing outcomes in high-pressure tasks.

The findings of this study suggest that while L2 learners exhibited higher anxiety levels than L1 learners, writing anxiety was a common challenge for both groups. Although no significant interaction between anxiety and task complexity was found, anxiety still seems to play a role in writing performance. Tasks with higher cognitive demands, such as argumentative writing, were linked to more linguistically complex outputs. This implies that the cognitive load associated with such tasks may heighten anxiety, which then negatively affects L2 writing. In conclusion, from a pedagogical perspective, these results highlight the importance of considering the interplay between anxiety, task complexity, and writing performance when designing writing tasks. Teachers should be particularly mindful of how cognitively demanding tasks, such as argumentative writing, can exacerbate anxiety, potentially hindering students' ability to produce high-quality work. To mitigate this, instructors should implement scaffolding strategies that support students throughout the writing process. These may include breaking tasks into smaller, manageable steps, providing structured guidance, and offering timely feedback. Additionally, fostering a classroom environment that reduces anxiety by encouraging a growth mindset, where challenges are viewed as opportunities for improvement, can help alleviate students' stress. By addressing both the cognitive and emotional aspects of writing tasks, educators can better support their students in overcoming writing anxiety, ultimately enhancing their writing performance and overall development as writers.

References

Abdi Tabari, M. (2022). Investigating the interactions between L2 writing processes and products under different task planning time conditions. *Journal of Second Language Writing*, *55*, 100871. https://doi.org/10.1016/j.jslw.2022.100871

Abdi Tabari, M. (2023). Unpacking the effects of different lengths of pre-task planning time: L2 writing outcomes and learners' perceptions. *The Language Learning Journal*, *52*(6), 615-628. https://doi.org/10.1080/09571736.2023.2213237

Abdi Tabari, M., Abdel Latif, M. M. M., & Tian, Y. (2024). The impact of computer-mediated task complexity on writing fluency: A comparative study of L1 and L2 writers' fluency performance. *Computers and Composition*, *73*, 102863. https://doi.org/10.1016/j.compcom.2024.102863

Abdi Tabari, M., & Goetze, J. (2024). Exploring the impacts of task complexity, anxiety, and self-efficacy on L2 written production: Unraveling individual differences in TBLT. *International Journal of Applied Linguistics*, *34*(4), 1533-1555. https://doi.org/10.1111/ijal.12582

Abdi Tabari, M., & Lee, J. (2024). The effects of task complexity on linguistic complexity and perception in English native speakers' and L2 speakers' argumentative writing. *Language Awareness*, *34*(1), 119-144. https://doi.org/10.1080/09658416.2024.2375226

Abdi Tabari, M., Farahanynia, M., & Botes, E. (2025). Examining the predictive power of L2 writing anxiety on L2 writing performance in simple and complex tasks under task-readiness conditions. *Assessing Writing*, *63*, 100912. https://doi.org/10.1016/j.asw.2024.100912

Albert, A. (2021). *Stories students tell: Creativity and narrative task performance.* Akadémiai Kiadó.

Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research, 15*(2), 201-229. https://doi.org/10.1177/0265532210388717

Barber, P. (2015). *Applied cognitive psychology: An information-processing framework*. Routledge.

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, *22*, 185-200. https://doi.org/10.1007/s11145-007-9107-5

Bi, P. (2020). Revisiting genre effects on linguistic features of L2 writing: A usage-based perspective. *International Journal of Applied Linguistics*, *30*(3), 429-444. https://doi.org/10.1111/ijal.12297

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1-28. https://doi.org/10.18637/jss.v080.i01

Busse, V., Graham, S., Müller, N., & Utesch, T. (2023). Understanding the interplay between text quality, writing self-efficacy, and writing anxiety in learners with and without migration background. *Frontiers in Psychology, 14*, 1130149. https://doi.org/10.3389/fpsyg.2023.1130149

Cakici, 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

Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning 25*, 153-161. https://doi.org/10.1111/j.1467-1770.1975.tb00115.x

Cheng, Y. (2002). Factors associated with foreign language writing anxiety. *Foreign Language Annals, 35*(6), 647-656. https://doi.org/10.1111/j.1944-9720.2002.tb01903.x

Cheng, Y., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417-446. https://doi.org/10.1111/0023-8333.00095

Cheng, Y. S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing, 13*(4), 313-335. https://doi.org/10.1016/j.jslw.2004.07.001

Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, *6*(3), 399-403.

Côté, S., & Gaffney, C. (2021). The effect of synchronous computer-mediated communication on beginner L2 learners' foreign language anxiety and participation. *The Language Learning Journal, 49*(1), 105-116. https://doi.org/10.1080/09571736.2018.1484935

Council of Europe (2020). *Common European framework of reference for languages: learning, teaching, assessment: Companion volume*. Council of Europe Publishing. https://www.coe.int/en/web/common-european-framework-reference-languages

Council of Europe (CEFR) (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge University Press.

Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, *9*(3), 242-249. https://doi.org/10.58680/rte197520067

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*(3), 193-220. https://doi.org/10.1177/136216880000400302

Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning*, *59*(4), 866-896. https://doi.org/10.1111/j.1467-9922.2009.00528.x

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457-472. https://doi.org/10.1214/ss/1177011136

Gkonou, C. (2011). Anxiety over EFL speaking and writing: A view from language classrooms. *Studies in Second Language Learning and Teaching, 1*(2), 267-281. https://doi.org/10.14746/ssllt.2011.1.2.6

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments & Computers, 36*(2), 193-202. https://doi.org/10.3758/BF03195564

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research, 15*(1), 1593-1623.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70*(2), 125-132. https://doi.org/10.2307/327317

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, *37*, 13-38. https://doi.org/10.1016/j.jslw.2017.06.001

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57-71). Erlbaum.

Kim, K. J., & Pae, T. I. (2021). Examining the simultaneous effects of L1 writing, L2 reading, L2 proficiency, and affective factors on different task types of L2 writing. *Discourse Processes, 58*(7), 662-680. https://doi.org/10.1080/0163853x.2021.1872989

Kim, Y.-S. G. (2020). Structural relations of language and cognitive skills, and topic knowledge to written composition: a test of the direct and indirect effects model of writing. *British Journal of Educational Psychology, 90,* 910-932. https://doi.org/10.1111/bjep.12330

Kleinmann, H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning, 27*, 93-107. https://doi.org/10.1111/j.1467-1770.1977.tb00294.x

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, *20*(2), 148-161. https://doi.org/10.1016/j.jslw.2011.02.001

Lambert, C., Aubrey, S., & Bui, G. (Eds.). (2023). *The role of the learner in task-based language teaching: Theory and research methods* (1st ed.). Routledge.

Lee, S. Y. (2007). Bayesian estimation of structural equation models. In S. Y. Lee (Ed.), *Structural equation modeling: A Bayesian approach* (pp. 67-109). John Wiley & Sons Inc.

Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 171-197). John Benjamins.

Li, C., Wei, L., & Lu, X. (2023). Contributions of foreign language writing emotions to writing achievement. *System, 116*, 103074. https://doi.org/10.1016/j.system.2023.103074

Li, S., & Yu, H. (2023). Effects of topic familiarity on L2 writing processes and behaviors. *International Journal of Applied Linguistics*, *34*(1), 348-366. https://doi.org/10.1111/ijal.12497

Long, M. (2014). *Second language acquisition and task-based language teaching*. Wiley.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*, 36-62. https://doi.org/10.5054/tq.2011.240859

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11-30). Multilingual Matters.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3-15. https://doi.org/10.1016/j.jslw.2015.06.004

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In S. Brooks, A. Gelman, G. Jones, & X.-L. Meng (Eds.), *Handbook of Markov chain Monte Carlo* (Vol. 2, pp. 113-162). Chapman and Hall/CRC.

R Core Team. (2022). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* In http://www.R-project.org/

Rahimi, M. (2018). Effects of increasing the degree of reasoning and the number of elements on L2 argumentative writing. *Language Teaching Research*, *23*(5), 633-654. https://doi.org/10.1177/1362168818761465

Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208-241. https://doi.org/10.1111/lang.12205

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, *21*(2), 45-105.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43,* 1-32. https://doi.org/10.1515/iral.2005.43.1.1

Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and directions in TBLT: Plenaries from a decade of the international conference* (Vol. 8, pp. 87-122). John Benjamins.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 167-185). Pearson Education.

Skehan, P. (2014). *Processing perspectives on task performance*. John Benjamins.

Stan Development Team. (2023). *Stan modeling language users guide and reference manual.* In Stan Development Team. https://mc-stan.org

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 495-508). Routledge.

Thompson, A. S., & Lee, J. (2014). The impact of experience abroad and language proficiency on language learning anxiety. *TESOL Quarterly*, *48*(2), 252-274. https://doi.org/10.1002/tesq.125

Trebits, A. (2016). Sources of individual differences in L2 narrative production: The contribution of input, processing, and output anxiety. *Applied Linguistics, 37*(2), 155-174. https://doi.org/10.1093/applin/amu006

Vytal, K. E., Cornwell, B. R., Letkiewicz, A. M., Arkin, N. E., & Grillon, C. (2013). The complex interaction between anxiety and cognition: insight from spatial and verbal working memory. *Frontiers in Human Neuroscience*, *7*, 93. https://doi.org/10.3389/fnhum.2013.00093

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Yang, W. (2014). *Mapping the relationships among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy, and fluency of L2 writing* [Unpublished PhD thesis]. Georgia State University.

Zabihi, R., Mousavi, S. H., & Salehian, A. (2020). The differential role of domain-specific anxiety in learners' narrative and argumentative L2 written task performances. *Current Psychology, 39*, 1438-1444. https://doi.org/10.1007/s12144-018-9850-6