Curriculum Research

Volume 7, Issue 2
Jun. 15, 2026

# Task Type and Task Complexity in Digitally Mediated L2 Writing: Effects on Syntactic Complexity, Accuracy, and Fluency

**Article Type:**

**Original Research**

**Authors:**

**Reyhaneh Daroudi**[1]

ORCiD: 0009-0001-9682-878X

**Mojgan Rashtchi**[2]

ORCiD: 0000-0001-7713-9316

**Parviz Maftoon**[3]

ORCiD: 0000-0002-5585-3647

**Article History:**

**Received:** 2026.01.29

**Accepted:** 2026.06.05

**Published:** 2026.06.15

## Abstract

This study investigated the independent and interactive effects of task type and task complexity on English as a Foreign Language (EFL) learners' writing performance. Sixty-three intermediate EFL learners completed two writing tasks that varied in task type (narrative vs. decision-making) and task complexity (simple vs. complex). Task complexity was manipulated through reasoning demands using a within-subjects design, while task type functioned as a between-subjects factor. Task stimuli were presented as animated video clips, and participants engaged in pre-task planning by annotating the videos. Writing performance was assessed using seven measures of Complexity, Accuracy, and Fluency (CAF): Mean Length of T-unit (MLT), Dependent Clauses per T-unit (DC/T), Complex Nominals per T-unit (CN/T), Complex Nominals per Clause (CN/C), Error-Free Ratio (EFR), Correct Verb Forms (CVF), and Text Length (TXL). Repeated-measures Multivariate Analysis of Variance (MANOVA) revealed significant effects of task type, task complexity, and their interaction. Decision-making tasks produced significantly higher MLT, CN/T, CN/C, EFR, and CVF than narrative tasks, whereas narrative tasks yielded greater fluency. Increasing task complexity increased MLT, CN/T, and CN/C but reduced accuracy, with no significant effects on DC/T or fluency. Significant Task Type × Task Complexity interactions were observed for MLT, CN/T, and CN/C, indicating that the effects of task complexity on syntactic complexity varied across task types. The findings demonstrate that task design exerts selective rather than uniform effects on L2 writing performance and that different dimensions of CAF respond differently to task manipulations.

**Key Words:** L2 writing, Task complexity, Task type, Writing Accuracy, Writing Complexity, Writing Fluency

1. Department of English, SR.C., Islamic Azad University, Tehran, Iran. Email: R.daroudi@iau.ir
2. Department of English, NT.C., Islamic Azad University, Tehran, Iran (Corresponding Author). Email: Mojgan.rashtchi@gmail.com
3. Department of English, SR.C., Islamic Azad University, Tehran, Iran. Email: pmaftoon@gmail.com

## 1. Introduction

Task-based Language Teaching (TBLT) has become one of the most influential approaches to second language (L2) instruction because it promotes language learning by engaging learners in meaning-focused tasks while encouraging attention to relevant linguistic forms (Ellis, 2003b). As TBLT has matured, however, a fundamental question has remained unresolved: which task characteristics most effectively facilitate L2 development? This question is central to both theory and pedagogy because task design can directly influence the cognitive and linguistic demands learners encounter during language production.

Among these characteristics, task complexity has received sustained theoretical and empirical attention. Two major accounts dominate this line of research. Robinson's Cognition Hypothesis (2001, 2003, 2005) argues that increasing cognitive complexity directs learners' attention to form, thereby facilitating simultaneous improvements in linguistic Complexity, Accuracy, and Fluency (CAF). In contrast, Skehan's Trade-off Hypothesis (1998, 2015) proposes that learners' limited attentional resources lead to trade-offs among CAF dimensions as task demands increase. These contrasting perspectives have made task complexity one of the most actively debated issues in task-based language research.

Although both frameworks have empirical support, findings remain inconsistent. While some studies report that increasing task complexity enhances linguistic complexity and, in some cases, accuracy (Kuiken & Vedder, 2007; Tabari & Lee, 2025), others document declines in accuracy or fluency or report no significant effects (Rashtchi & Mohammad Yousefi, 2017). These inconsistencies suggest that the effects of task complexity may not operate independently but instead depend on other task design characteristics.

One task characteristic that may explain these divergent findings is task type. Different writing tasks impose distinct communicative purposes, discourse structures, and cognitive operations, which may alter how learners allocate their limited attentional resources (Foster & Skehan, 1996; Kuiken & Vedder, 2008). For example, narrative tasks typically require learners to construct temporally organized discourse, whereas decision-making tasks emphasize evaluating alternatives, weighing evidence, and justifying choices. These differences imply that identical complexity manipulations may yield different linguistic outcomes. Nevertheless, despite the theoretical importance of this possibility, relatively little research has examined whether task type moderates the effects of task complexity on L2 writing performance (Allaw, 2019; Gilabert et al., 2009).

The limited evidence available suggests that the interaction between task type and task complexity warrants closer investigation. Existing studies have generally examined these variables separately or focused predominantly on oral production, leaving important questions about written language unresolved (Bui et al., 2018; Kim, 2023; Pang & Skehan, 2021; Robinson, 2005; Skehan & Foster, 2001). Consequently, it remains unclear whether inconsistencies reported in previous research reflect genuine limitations of current theoretical models or arise because the interaction between task characteristics has not been adequately considered. Examining task type and task complexity simultaneously therefore offers an opportunity not merely to compare two task variables but to clarify the conditions under which competing predictions derived from the Cognition Hypothesis and the Trade-off Hypothesis are supported.

The present study addressed this gap by investigating the independent and interactive effects of task type (narrative vs. decision-making) and task complexity (simple vs. complex) on EFL learners' writing performance across CAF. Guided by the Cognition Hypothesis, task complexity was operationalized as the reasoning demands inherent in task design; that is, the quantity and intensity of cognitive operations required to complete a task (Robinson, 2001). The study further sought to extend previous research by examining these task characteristics in a digitally mediated writing environment with multimedia task presentation, thereby reflecting contemporary technology-enhanced language learning contexts. Beyond its empirical contribution, this study sought to advance current theoretical understanding by examining whether the effects of reasoning-demand complexity depend on the task's communicative demands. From a pedagogical perspective, the findings may inform the design of cognitively appropriate writing tasks for digitally mediated language classrooms, enabling teachers, curriculum developers, and materials designers to sequence tasks more effectively in line with learners' cognitive and linguistic needs.

## 2. Review of the Related Literature

Task type has emerged as a determinant of L2 performance because different tasks impose distinct communicative purposes, discourse structures, and cognitive operations. Rather than serving merely as alternative elicitation formats, task types direct learners' attentional resources toward different aspects of language production, shaping the linguistic features that emerge in written output (Foster & Skehan, 1996; Gilabert et al., 2009; Skehan, 2016). Task classifications proposed within TBLT largely converge on distinguishing narrative and decision-making tasks as cognitively and rhetorically distinct genres. Narrative tasks primarily require learners to construct coherent, temporally organized discourse, whereas decision-making tasks involve evaluating alternatives, integrating information, and justifying choices (Foster & Skehan, 1996; Gilabert et al., 2009).

Research generally supports the influence of task type on L2 writing performance, though the effects vary across linguistic dimensions. Studies comparing narrative, decision-making, argumentative, descriptive, integrated, and independent writing tasks have consistently shown that different task types elicit distinct linguistic profiles (Gonzalez & Mayo, 2025; Jiang et al., 2025; Liu et al., 2025; Yoon, 2021). Narrative tasks are often associated with greater lexical and discourse development, whereas decision-making and integrated tasks often promote greater syntactic elaboration and analytical language use (Garcia-Ponce & Tavakoli, 2022; Hsieh, 2024; Zhao et al., 2024).

Evidence also suggests that task type affects CAF dimensions differently rather than uniformly. Narrative and descriptive tasks have been found to promote lexical richness, discourse elaboration, and certain measures of accuracy, whereas decision-making, argumentative, and integrated tasks often facilitate higher syntactic complexity but, in some cases, lower fluency (Bao, 2025; Gonzalez & Mayo, 2025; Zhao et al., 2025).

Similar task-dependent effects have also been observed in oral production, where storytelling, opinion exchange, and descriptive tasks yielded distinct patterns of lexical complexity, interaction, and accuracy (Gilanlioglu, 2022; Qiu & Cheng, 2022). Meanwhile, the effects of task type appear to depend partly on learner- and context-related variables, including planning, proficiency, working memory, and instructional support (Garcia-Ponce & Tavakoli, 2022; Jiang et al., 2025; Zhao et al., 2025). These findings suggest that task type rarely operates in isolation but interacts with other factors that influence language production.

Despite broad agreement that task type influences L2 writing performance, important inconsistencies remain. While some studies report advantages for cognitively demanding task types in syntactic complexity and linguistic sophistication, others report reduced fluency, lower accuracy, or negligible differences across CAF measures (Diez-Bedmar & Thewissen, 2025; Gonzalez & Mayo, 2025; Zhao et al., 2025). Such variability suggests that task type effects are likely conditioned by other task characteristics, particularly task complexity. Examining these variables jointly may therefore provide a more comprehensive explanation for the mixed findings reported in previous TBLT research.

Task complexity has long been regarded as a central construct in TBLT because it determines the cognitive demands learners must manage during task performance. Despite extensive research, its effects on L2 writing remain inconclusive. This lack of consensus has generally been interpreted through two competing theoretical perspectives. Skehan's (1998, 2015) Limited Attentional Capacity Model proposes that learners possess finite cognitive resources and therefore cannot simultaneously prioritize all aspects of language production. As task demands increase, attention must be distributed selectively, leading to trade-offs among CAF dimensions.

However, Robinson's (2001, 2003, 2007, 2015) Cognition Hypothesis argues that increasing task complexity can channel learners' attention toward linguistically relevant features, promoting greater linguistic complexity without necessarily compromising accuracy or fluency. Within Robinson's Triadic Componential Framework (2003), task complexity is distinguished from task difficulty and task conditions. Task complexity refers to the cognitive demands inherent in task design, whereas task difficulty reflects learner characteristics, and task conditions concern interactional variables. A key distinction is between resource-directing and resource-dispersing variables. Resource-directing variables, such as reasoning demands, are expected to encourage greater conceptual processing and linguistic elaboration, whereas resource-dispersing variables primarily increase processing load. The present study focused on reasoning demands, one of the most frequently investigated resource-directing dimensions.

Empirical evidence supports both theoretical accounts. Some studies have shown that increasing reasoning demands promotes greater syntactic complexity and lexical sophistication, consistent with the Cognition Hypothesis (Robinson, 2007; Tabari & Lee, 2025; Zhan et al., 2021). Others have reported simultaneous improvements across multiple CAF dimensions when cognitively demanding tasks were accompanied by instructional support, such as strategic planning or task sequencing (Ryu & Lee, 2018; Tabari & Miller, 2021). Conversely, some studies have found trade-offs: increased task complexity enhances some aspects of performance while reducing others. For example, greater reasoning demands have been associated with increased syntactic complexity but lower accuracy, lexical sophistication, or fluency (Abrams, 2019; Rahimi & Zhang, 2019; Tao & Wang, 2025), whereas other studies have reported weak or non-significant effects across CAF measures (Xu et al., 2023).

The coexistence of these contrasting findings suggests that the effects of task complexity may be conditional rather than universal, and that neither theoretical framework consistently explains learners' performance across task contexts. Recent research increasingly indicates that learner characteristics, planning opportunities, and task design may moderate learners' responses to cognitively demanding tasks (Jin & Yan, 2025; Tao & Wang, 2025). Nevertheless, comparatively little attention has been paid to whether the influence of reasoning-demand complexity depends on the communicative demands imposed by different task types. Examining this interaction may help explain the inconsistent findings reported in previous research and clarify the conditions under which the competing predictions of the Cognition Hypothesis and the Limited Attentional Capacity Model are supported.

Beyond investigating individual task variables, several studies have explored how multiple aspects of task design operate together during task performance. These studies provide evidence that the effects of one task characteristic may depend on the presence of another. For example, Gilabert (2007) demonstrated that increasing reasoning demands produced different effects across task types, suggesting that identical complexity manipulations do not necessarily yield comparable linguistic outcomes across communicative contexts. Similarly, McDonough and Fuentes (2015) found that writing accuracy was jointly influenced by task type and task implementation conditions rather than by either variable alone. Subsequent studies have reinforced this view. Alexopoulou et al. (2017) reported that the effects of cognitive complexity varied across task types. Allaw (2019) showed that task complexity interacted with task type to influence propositional and linguistic complexity. Lee (2020) also found that task type moderated the relationship between task complexity, lexical diversity, and accuracy.

Collectively, these studies suggest that the effects of task demands depend on the communicative goals and discourse structures of different task types, highlighting the need to examine task characteristics as interrelated variables. Despite increasing attention to the role of task characteristics, important gaps remain. Most previous studies have focused on oral production (Bui et al., 2018; Gilabert, 2007; Pang & Skehan, 2021), examined a limited range of linguistic measures (Gonzalez & Mayo, 2025), or investigated traditional instructional settings (Alexopoulou et al., 2017; Lee, 2020). Comparatively little research has explored whether reasoning-demand task complexity interacts with different writing task types to influence multiple measures of CAF, particularly in multimedia-supported learning environments.

The present study addressed this gap by examining the independent and interactive effects of task type and task complexity on EFL learners' writing performance. By investigating whether complexity effects differ between narrative and decision-making tasks, the study sought to clarify the conditions under which competing predictions derived from the Cognition Hypothesis and the Limited Attentional Capacity Model are supported, while extending task-based writing research to technology-enhanced instructional contexts. Accordingly, the study addressed the following research questions:

**RQ1.** What are the individual effects of task type and task complexity on EFL learners' writing performance in terms of syntactic complexity, accuracy, and fluency?

**RQ2.** What is the interaction effect of task type and task complexity on EFL learners' writing performance in terms of syntactic complexity, accuracy, and fluency?

## 3. Method

### 3.1. Research Design

The study employed a 2 × 2 mixed-factorial cluster quasi-experimental design to examine the effects of task type and task complexity on L2 written performance. Task type (narrative vs. decision-making) served as the between-subjects factor and was assigned at the level of intact classes rather than individual participants, whereas task complexity (simple vs. complex) constituted the within-subjects factor. Seven CAF measures served as dependent variables. Syntactic complexity was assessed using Mean Length of T-unit (MLT), Dependent Clauses per T-unit (DC/T), Complex Nominals per T-unit (CN/T), and Complex Nominals per Clause (CN/C). Accuracy was measured using the Error-Free Ratio (EFR) and Correct Verb Forms (CVF), while fluency was operationalized as total Text Length (TXL) produced within the fixed writing time. Participants were screened for comparable English proficiency before the experiment. The sample consisted exclusively of female learners.

### 3.2. Participants

Participants were selected from intermediate-level English classes at a private language school. The initial sample comprised 79 female EFL learners who were non-English majors, aged 18 to 31. English proficiency was assessed using the Preliminary English Test (PET; Cambridge Assessment English). Following recommendations commonly adopted in experimental SLA research, only learners whose scores fell within ±1.5 standard deviations of the sample mean were retained, yielding 71 eligible participants. The participants were assigned to task-type conditions at the class level rather than through individual random assignment. During data collection, eight participants withdrew due to absenteeism, resulting in a final sample of 63 learners. Of these, 28 participants completed the narrative writing tasks, and 35 participants completed the decision-making tasks.

## 3.3. Instruments

### *3.3.1. Proficiency Measure*

English proficiency was assessed using the PET. The test evaluates reading, writing, listening, and speaking skills at the intermediate level. Two EFL teachers independently evaluated the productive components using the PET assessment criteria. The inter-rater reliability analysis indicated good agreement between the raters (ICC = .87), and the ratings were subsequently averaged to obtain the final writing and speaking scores. The PET scores were used as a screening measure to ensure a relatively homogeneous proficiency level among participants. Internal consistency for the present sample was satisfactory (Cronbach's $\alpha = .82$).

### *3.3.2. Writing Tasks*

Written performance was elicited through four writing tasks representing two task types (narrative vs. decision-making), each operationalized at two levels of cognitive complexity (simple vs. complex). Task complexity was manipulated by varying reasoning demands while maintaining the same communicative goal across task types.

The narrative tasks were adapted from the WAIS-III Picture Arrangement subtest (stories 5 and 11). Although the WAIS-III is formally an intelligence assessment, its Picture Arrangement subtest has been adopted in SLA research as a narrative elicitation task (Révész et al., 2016; Robinson, 2007). The simple version depicted a routine sequence of events that required minimal inference, whereas the complex version required interpreting characters' intentions and integrating multiple perspectives. To reduce extraneous cognitive demands, picture sequencing was removed, and all images were presented in a predetermined order.

Decision-making tasks were adapted from the Fire Chief Task (Gilabert et al., 2009), in which participants evaluated a simulated emergency scenario and justified their rescue decisions. The simple condition involved uniform victims and stable conditions, while the complex condition introduced heterogeneous victims, dynamic hazards, and limited rescue resources, requiring prioritization and higher-order reasoning.

All task materials were presented as short animated video clips to standardize task presentation and reduce variability in stimulus delivery. Minor visual modifications were made to improve clarity while preserving the conceptual structure of the original tasks. Before each task, participants received written instructions. For the narrative tasks, they were instructed to write a story based on the video. For the decision-making tasks, they were instructed to describe and justify their decisions in response to the emergency scenario. Participants were given 10 minutes for pre-task planning, during which they could use the platform's note-taking function, followed by 30 minutes to complete the writing task. They were not permitted to use external sources and were encouraged to include as many relevant details as possible in their written responses.

#### *3.3.3. Validation of the Task Complexity Manipulation*

The task complexity manipulation was established through multiple sources of evidence. First, the task designs were grounded in prior empirical studies that demonstrated differences in cognitive demands between the simple and complex versions (Gilabert, 2007; Révész et al., 2016; Vasylets et al., 2017). Second, two Applied Linguistics experts reviewed the adapted tasks to assess their appropriateness for the complexity manipulation. Third, the manipulation was empirically verified through a pilot study with 16 learners comparable to those in the main study. The results consistently indicated that the complex versions imposed greater cognitive demands.

### 3.4. Procedure

#### 3.4.1. Pilot Studies

Two pilot studies were conducted prior to the main experiment. The first pilot established appropriate planning and writing durations. Five intermediate EFL learners completed all experimental tasks without time restrictions. Their completion times, together with recommendations from previous task-planning research (Ellis, 2003a; Foster & Skehan, 1996; Rahimi & Zhang, 2018; Skehan & Foster, 2005), informed the allocation of a 10-minute planning phase followed by a 30-minute writing phase.

The second pilot study evaluated the effectiveness of the task complexity manipulation using a separate sample of 16 learners. Participants completed post-task ratings of perceived task difficulty (Révész et al., 2016) and mental effort (Malicka, 2020), confirming that the complex task versions were consistently perceived as more cognitively demanding than the corresponding simple versions.

#### 3.4.2. Main Study

The experiment was conducted with intact classes, randomly assigned at the class level to one of two task-type conditions (narrative vs. decision-making). Four intact classes were assigned to each task-type condition. All classes were taught by the same instructor using identical instructional procedures, materials, and testing conditions to minimize classroom-level variability. Because only eight classes participated in the study, multilevel modeling was not considered appropriate; therefore, all statistical analyses were conducted at the individual level.

An Independent-samples *t*-test on participants' PET scores indicated no significant differences between groups, suggesting that they were comparable in English language proficiency before the experimental treatments. Next, the eligible participants completed two writing tasks within their assigned task type: a simple and a complex version. To reduce potential fatigue and practice effects, the two tasks were administered across two sessions separated by no more than one week. Within each intact class, the order of the simple and complex tasks was counterbalanced so that approximately half of the participants completed the simple task first and the remainder completed the complex task first.

At the beginning of each session, participants received standardized written instructions and viewed the animated task stimuli during a 10-minute planning period. They were permitted to pause and annotate the videos using the ReClipped platform, which enabled participants to record observations while ensuring identical access to the task materials. To prevent planning notes from influencing online language production, these annotations were inaccessible during the subsequent writing phase.

Following planning, participants completed the writing task individually within 30 minutes under standardized testing conditions without access to external resources. All experimental sessions were conducted using identical procedures and time allocations to maximize procedural consistency across groups. To preserve ecological validity, the experiment was integrated into regularly scheduled classroom instruction.

Permission to conduct the study was obtained from the administration of the participating language school. At the time the study was conducted, the institution did not require formal ethical approval for educational research of this nature. Nevertheless, the study adhered to internationally accepted ethical principles for research involving human participants. Prior to participation, all learners provided informed consent and were informed that their involvement was voluntary and that withdrawal could occur at any time without academic consequences. They were assured that all responses would remain confidential, would be used exclusively for research purposes, and would not influence their course grades. The specific research objectives were withheld until data collection was complete to minimize expectancy effects.

### 3.4.3. CAF Measures

The selection of CAF measures was guided by four criteria: (a) the inclusion of general measures considered sensitive to group differences in experimental research (Skehan & Foster, 1999); (b) the inclusion of specific measures recommended by Norris and Ortega (2009); (c) appropriateness for written production; and (d) consistency with measures used in previous studies (Foster & Skehan, 1996; Kuiken & Vedder, 2007; Robinson, 2007).

Seven indices were selected to represent the three CAF dimensions. Syntactic complexity was operationalized using four indices. MLT captured overall syntactic elaboration, DC/T indexed subordination, while CN/T and CN/C measured phrasal elaboration.

Accuracy was assessed using two complementary measures. EFR, as a general measure, was calculated as the percentage of error-free clauses relative to the total number of clauses produced. CVF, as a specific measure, assessed the proportion of verbs accurately produced with respect to tense, aspect, modality, and agreement. Error identification followed Polio's (1997) coding framework, except that spelling-related features, including punctuation and capitalization, were treated separately from the original coding scheme.

Fluency was operationalized as total text length (TXL), calculated as the total number of words produced during the fixed 30-minute writing period. Hyphenated forms were counted as a single lexical unit following Gilabert (2005).

All writing samples were coded according to the predefined CAF criteria. Two trained raters with graduate-level training in applied linguistics independently coded all texts following a detailed coding manual and a series of calibration sessions using a subset of pilot scripts. Prior to the main coding phase, the raters jointly coded 10% of the data to establish consistency and clarify ambiguous cases. Inter-rater reliability was high across the CAF measures, with Cohen's κ ranging from .84 to .92 for categorical accuracy codes and intraclass correlation coefficients (ICC) ranging from .88 to .95 for continuous complexity and fluency indices. Any discrepancies were resolved through discussion, and when consensus could not be reached, a third experienced coder adjudicated the final decision.

### 3.4.4. Data Analysis

Data were analyzed using SPSS Version 27. Before the main analyses, the assumptions underlying multivariate analysis were evaluated. A repeated-measures Multivariate Analysis of Variance (MANOVA) was then conducted to examine the main effects of task type and task complexity, as well as their interaction, across seven correlated measures of syntactic complexity, accuracy, and fluency. MANOVA was selected because it simultaneously analyzes multiple correlated dependent variables while controlling the familywise Type I error rate. Consistent with the study's a priori analytic plan, significant multivariate effects were followed by separate analyses for each CAF dimension to identify the specific measures contributing to the overall effects. Statistical significance was set at $\alpha = .05$, and partial eta squared ($\eta p^2$) was reported as the measure of effect size.

## 4. Results

### 4.1. Preliminary Analyses

PET scores for the initial sample (N = 79) were normally distributed, $W(79) = .98$, $p = .25$. Applying the predetermined screening criterion (±1.5 SD) yielded a final sample of 71 participants. An independent-samples *t* test confirmed no significant difference in PET scores between the narrative and decision-making groups, $t(69) = 0.18$, $p = .85$. Assumptions for the repeated-measures MANOVA, including normality (boxplots and Shapiro–Wilk tests), homogeneity of variance (Levene's tests), and equality of covariance matrices (Box's *M* test), were all met ($p > .05$). Because task complexity had only two levels, sphericity was assumed.

### 4.2. Effects of Task Type and Task Complexity on CAF

As shown in Table 1, the repeated-measures MANOVA revealed significant main effects of task type (Pillai's Trace = .806, $F_{(7, 55)} = 32.58$, $p < .001$, $\eta p^2 = .80$) and task complexity (Pillai's Trace = .831, $F_{(7, 55)} = 38.53$, $p < .001$, $\eta p^2 = .83$), as well as a significant interaction (Pillai's Trace = .671, $F_{(7, 55)} = 16.00$, $p < .001$, $\eta p^2 = .67$).

**Table 1.**

*Multivariate Tests of the Effects of Task Type, Task Complexity, and Their Interaction on L2 Writing Performance*

| Effect | Pillai's Trace | $F(7, 55)$ | $P$ | $\eta p^2$ |
|---|---|---|---|---|
| Task Type | .806 | 32.58 | < .001 | .80 |
| Task Complexity | .831 | 38.53 | < .001 | .83 |
| Task Complexity * Task Type | .671 | 16.00 | < .001 | .67 |

To determine which dependent variables contributed to these multivariate effects, follow-up univariate repeated-measures analyses were conducted. Descriptive statistics are presented in Table 2.

**Table 2.**

*Descriptive Statistics for CAF Measures by Task Type, Task Complexity, and Their Interaction*

| | Task Type | | | Task Complexity | | | Interaction | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Measure | Task Type | M | SD | Task Complexity | M | SD | Task Type | Task Complexity | M | SE |
| MLT | Na | 13.18 | 1.39 | Simple | 13.72 | 2.28 | Na | Simple | 12.44 | .37 |
| | | | | | | | | Complex | 13.91 | .39 |
| | De | 16.24 | 1.45 | Complex | 16.04 | 2.81 | De | Simple | 14.74 | .33 |
| | | | | | | | | Complex | 17.74 | .35 |
| DC/T | Na | .72 | .11 | Simple | .72 | .16 | Na | Simple | .74 | .03 |
| | | | | | | | | Complex | .70 | .03 |
| | De | .69 | .13 | Complex | .69 | .15 | De | Simple | .70 | .02 |
| | | | | | | | | Complex | .68 | .02 |
| CN/T | Na | .99 | .23 | Simple | 1.07 | .33 | Na | Simple | .91 | .05 |
| | | | | | | | | Complex | 1.08 | .05 |
| | De | 1.39 | .22 | Complex | 1.36 | .39 | De | Simple | 1.20 | .05 |
| | | | | | | | | Complex | 1.59 | .05 |
| CN/C | Na | .56 | .17 | Simple | .61 | .32 | Na | Simple | .37 | .04 |
| | | | | | | | | Complex | .76 | .05 |
| | De | .90 | .20 | Complex | .88 | .28 | De | Simple | .81 | .04 |
| | | | | | | | | Complex | .99 | .04 |
| EFR | Na | .62 | .06 | Simple | .72 | .12 | Na | Simple | .68 | .02 |
| | | | | | | | | Complex | .56 | .02 |
| | De | .66 | .07 | Complex | .57 | .10 | De | Simple | .74 | .02 |
| | | | | | | | | Complex | .58 | .01 |
| CVF | Na | .74 | .05 | Simple | .80 | .08 | Na | Simple | .78 | .01 |
| | | | | | | | | Complex | .69 | .01 |
| | De | .78 | .04 | Complex | .72 | .07 | De | Simple | .81 | .01 |
| | | | | | | | | Complex | .75 | .01 |
| TXL | Na | 313.53 | 58.54 | Simple | 299.78 | 81.68 | Na | Simple | 316.96 | 15.28 |
| | | | | | | | | Complex | 310.10 | 13.28 |
| | De | 285.25 | 51.24 | Complex | 295.87 | 70.90 | De | Simple | 286.02 | 13.66 |
| | | | | | | | | Complex | 284.48 | 11.88 |

*Note.* Na = Narrative; De = Decision-making

The univariate analyses indicated significant Task Type × Task Complexity interactions for MLT, CN/T, and CN/C. Accordingly, the interpretation of the corresponding main effects should be made with caution and in conjunction with the simple-effects analyses reported later in this section. For the remaining dependent variables (DC/T, EFR, CVF, and TXL), the interaction effects were not statistically significant; therefore, the corresponding main effects can be interpreted directly (Table 3).

**Table 3.**

*Univariate Tests of the Effects of Task Type, Task Complexity, and Their Interaction on Individual CAF Measures*

| | Task Type | | | Task Complexity | | | Interaction | | |
|---|---|---|---|---|---|---|---|---|---|
| **Measure** | **F (1, 61)** | ***P*** | **ηp²** | **F (1, 62)** | ***P*** | **ηp²** | **F (1, 61)** | ***P*** | **ηp²** |
| **MLT** | 71.31 | < .001 | .54 | 38.05 | < .001 | .38 | 4.32 | .042 | .06 |
| **DC/T** | 1.01 | .320 | .02 | 1.53 | .220 | .02 | .06 | .793 | .00 |
| **CN/T** | 47.48 | < .001 | .44 | 32.90 | < .001 | .34 | 4.83 | .032 | .07 |
| **CN/C** | 48.49 | < .001 | .44 | 39.23 | < .001 | .38 | 6.15 | .016 | .09 |
| **EFR** | 5.19 | .026 | .07 | 44.97 | < .001 | .42 | .88 | .351 | .01 |
| **CVF** | 13.67 | < .001 | .18 | 30.45 | < .001 | .32 | .41 | .524 | .00 |
| **TXL** | 4.17 | .045 | .06 | .08 | .767 | .00 | .04 | .843 | .00 |

### *4.2.1. Task Type Effects*

Because the multivariate analysis revealed a significant Task Type × Task Complexity interaction, the main effects of task type are reported as overall effects and should be interpreted in conjunction with the interaction analyses.

#### *4.2.1.1. Syntactic Complexity*

At the multivariate level, task type significantly affected the combined syntactic complexity measures, Pillai's Trace = .760, $F_{(4,\ 58)}$ = 45.84, $p$ < .001, ηp² = .76. At the univariate level, significant effects of task type were observed for MLT, CN/T, and CN/C (Table 3). The mean MLT score was significantly higher for the decision-making task ($M$ = 16.24, $SD$ = 1.45) than for the narrative task ($M$ = 13.18, $SD$ = 1.39), $F_{(1,\ 61)}$ = 71.31, $p$ < .001, *ηp²* = .54. Similarly, CN/T was higher in the decision-making condition ($M$ = 1.39, $SD$ = .22) than in the narrative condition ($M$ = .99, $SD$ = .23), $F_{(1,\ 61)}$ = 47.48, $p$ < .001, ηp² = .44. A comparable pattern was observed for CN/C, with higher values in the decision-making condition ($M$= .90, $SD$ = .20) than in the narrative condition ($M$=.56; $SD$ = .17), $F_{(1,\ 61)}$ = 48.49, $p$ < .001, ηp² = .44. In contrast, although narrative tasks showed slightly higher levels of subordination than decision-making tasks ($M$ = .72, $SD$ = .11 vs. $M$ = .69, $SD$ = .13), the effect of task type on DC/T was not statistically significant, $F_{(1,\ 61)}$ = 1.01, $p$ = .320, ηp² = .02.

#### *4.2.1.2. Accuracy*

Task type also had a significant multivariate effect on the combined accuracy measures, Pillai's Trace = .470, $F_{(2, 60)}$ = 27.56, $p$ < .001, ηp² = .47. Tables 4.2 and 4.3 show that learners in the decision-making condition obtained higher EFR scores ($M$ = .66, $SD$ = .07) than those in the narrative condition ($M$ = .62, $SD$ = .06), $F_{(1, 61)}$ = 5.19, $p$ = .026, ηp² = .07. Likewise, CVF was significantly higher in the decision-making ($M$ = .78, $SD$ = .04) than in the narrative condition ($M$ = .74, $SD$ = .05), $F_{(1, 61)}$ = 13.67, $p$ < .001, ηp² = .18.

#### *4.2.1.3. Fluency*

Task type significantly affected text length, $F_{(1, 61)}$ = 4.17, $p$ = .045, ηp² = .06. Participants in the narrative condition produced longer texts ($M$ = 313.53, $SD$ = 58.54) than those in the decision-making condition ($M$ = 285.25, $SD$ = 51.24).

Overall, these findings suggest that learners performing the decision-making task obtained significantly higher scores on MLT, CN/T, CN/C, and accuracy. By contrast, learners performing the narrative task achieved significantly higher fluency scores. No statistically significant difference in DC/T between task types was observed. Because significant interactions were identified for MLT, CN/T, and CN/C, the main effects should be interpreted alongside the corresponding interaction analyses reported below.

### *4.2.2. Task Complexity Effects*

Because the multivariate analysis revealed a significant Task Type × Task Complexity interaction, the task complexity results below are reported as overall effects across task types and should be interpreted together with the interaction analyses.

#### *4.2.2.1. Syntactic Complexity*

Task complexity had a significant multivariate effect on the combined syntactic complexity measures, Pillai's Trace = .700, $F_{(4, 59)}$ = 34.88, $p$ < .001, ηp² = .70.

The descriptive and inferential results (Tables 2 and 3) show that learners produced significantly longer T-units in the complex condition ($M$ = 16.04, $SD$ = 2.81) than in the simple condition ($M$ = 13.72, $SD$ = 2.28), $F_{(1, 62)}$ = 38.05, $p$ < .001, ηp² = .38. Similarly, CN/T was significantly higher under the complex condition ($M$ = 1.36, $SD$ = 0.39) than under the simple condition ($M$ = 1.07, $SD$ = 0.33), $F_{(1, 62)}$ = 32.90, $p$ < .001, ηp² = .34. Participants also produced higher CN/C values in the complex condition ($M$ = 0.88, $SD$ = 0.28) than in the simple condition ($M$ = 0.61, $SD$ = 0.32), $F_{(1, 62)}$ = 39.23, $p$ < .001, ηp² = .38. In contrast, DC/T showed a slight decline from the simple (*M* = 0.72, *SD* = 0.16) to the complex condition (*M* = 0.69, *SD* = 0.15); however, the effect was not statistically significant, $F_{(1, 62)} = 1.53$, $p = .220$, ηp² = .02.

### *4.2.2.2. Accuracy*

Task complexity also significantly affected the combined accuracy measures, Pillai's Trace = .490, $F_{(2, 61)} = 29.38$, $p < .001$, ηp² = .49. Accuracy declined as task complexity increased. Learners produced significantly higher error-free ratios in the simple condition (*M* = 0.72, *SD* = 0.12) than under the complex condition (*M* = 0.57, *SD* = .10), $F_{(1, 62)} = 44.97$, $p < .001$, ηp² = .42. Similarly, CVF was higher in the simple condition (*M* = 0.80, *SD* = 0.08) than in the complex condition (*M* = 0.72, *SD* = 0.07), $F_{(1, 62)} = 30.45$, $p < .001$, ηp² = .32.

### *4.2.2.3. Fluency*

Fluency remained relatively stable across task complexity levels. Learners produced slightly more words under the simple (*M* = 299.78, *SD* = 81.68) than the complex task condition (*M* = 295.87, *SD* = 70.90); however, this difference was not statistically significant, $F_{(1, 62)} = .08$, $p = .767$, ηp² = .001.

These findings indicate that increased task complexity promoted greater syntactic elaboration (MLT, CN/T, CN/C) while simultaneously constraining writing accuracy, with no measurable effect on DC/T or fluency. These overall effects should be interpreted alongside the significant interaction between task type and task complexity.

### *4.2.3. Interaction Effects of Task Type and Task Complexity on CAF*

The significant multivariate interaction (Table 1) prompted follow-up univariate analyses to identify which dependent variables contributed to the effect. Significant interactions were observed for three syntactic complexity measures: MLT, $F(1, 59) = 4.32$, $p = .042$, ηp² = .06; CN/C, $F_{(1, 59)} = 6.15$, $p = .016$, ηp² = .07; CN/T, $F_{(1, 59)} = 4.83$, $p = .032$, ηp² = .09. In contrast, no significant interaction effects were found for DC/T, EFR, CVF, or TXL (*ps* > .05).

To interpret the significant interactions, simple-effects analyses were conducted within each task type, followed by pairwise comparisons between task types at each level of task complexity (Table 4). Estimated Marginal Means (EMMs) are presented to facilitate interpretation (Table 1).

**Table 4.**

*Simple Effects of Task Complexity and Pairwise Comparisons for Significant Task Type × Task Complexity Interactions*

| Measure | Analysis | Comparison | Statistic | *p* | Effect size |
|---|---|---|---|---|---|
| MLT | Simple effect | Narrative: Simple vs. Complex | $F(1, 27) = 13.80$ | .001 | ηp² = .33 |
| | | Decision-making: Simple vs. Complex | $F(1, 34) = 26.94$ | < .001 | ηp² = .44 |
| | Pairwise comparison | Simple: Narrative vs. Decision-making | $t(61) = -4.53$ | < .001 | $d = 1.15$ |
| | | Complex: Narrative vs. Decision-making | $t(61) = -7.25$ | < .001 | $d = 1.84$ |
| CN/T | Simple effect | Narrative: Simple vs. Complex | $F(1, 27) = 9.02$ | .006 | ηp² = .25 |
| | | Decision-making: Simple vs. Complex | $F(1, 34) = 25.96$ | < .001 | ηp² = .43 |
| | Pairwise comparison | Simple: Narrative vs. Decision-making | $t(59.90) = -4.03$ | < .001 | $d = 0.98$ |
| | | Complex: Narrative vs. Decision-making | $t(61) = -6.67$ | < .001 | $d = 1.70$ |
| CN/C | Simple effect | Narrative: Simple vs. Complex | $F(1, 27) = 31.24$ | < .001 | ηp² = .53 |
| | | Decision-making: Simple vs. Complex | $F(1, 34) = 12.54$ | .001 | ηp² = .27 |
| | Pairwise comparison | Simple: Narrative vs. Decision-making | $t(61) = -7.20$ | < .001 | $d = 1.83$ |
| | | Complex: Narrative vs. Decision-making | $t(61) = -3.49$ | .001 | $d = 0.89$ |

As shown in Table 4, task complexity significantly increased MLT scores in both the narrative task, $F_{(1, 27)} = 13.80$, $p = .001$, ηp² = .33, and the decision-making task, $F_{(1, 34)} = 26.94$, $p < .001$, ηp² = .44. The EMM increased from 12.44 to 13.91 in the narrative task (Δ = 1.47) and from 14.74 to 17.74 (Δ = 3.00) in the decision-making task, indicating a larger increase in the decision-making condition. Pairwise comparisons further showed that decision-making tasks produced significantly higher MLT than narrative tasks in both the simple ($t(61) = 4.53$, $p < .001$, d = 1.15) and complex ($t(61) = 7.25$, $p < .001$, d = 1.84) conditions.

A similar pattern was observed for CN/T. Task complexity significantly increased CN/T in both narrative, $F_{(1, 27)} = 9.02$, $p = .006$, $ηp^2 = .25$, and decision-making tasks, $F_{(1, 34)} = 25.96$, $p < .001$, $ηp^2 = .43$. The EMM increased from 0.91 to 1.08 in the narrative condition (Δ = 0.17) and from 1.20 to 1.59 in the decision-making condition (Δ = 0.38). Pairwise comparisons indicated that decision-making tasks produced significantly higher CN/T scores than narrative tasks under both the simple condition ($t(59.90) = 4.03$, $p < .001$, $d = 0.98$) and the complex condition ($t(61) = 6.67$, $p < .001$, $d = 1.70$).

For CN/C, task complexity significantly increased scores for both the narrative task, $F_{(1, 27)} = 31.24$, $p < .001$, $ηp^2 = .53$, and the decision-making task, $F_{(1, 34)} = 12.54$, $p = .001$, $ηp^2 = .27$. The EMM increased from 0.37 to 0.76 in the narrative task (Δ = 0.39) and from 0.81 to 0.99 in the decision-making task (Δ = 0.18), indicating a larger increase for the narrative task. Nevertheless, across task types, decision-making tasks produced significantly higher CN/C values than narrative tasks under both simple ($t(61) = 7.20$, $p < .001$, $d = 1.83$) and complex ($t(61) = 3.49$, $p = .001$, $d = 0.89$) conditions.

## 5. Discussion

The present study examined the effects of task type and task complexity on L2 written performance across syntactic complexity, accuracy, and fluency. Overall, both variables significantly influenced writing performance, although their effects differed across CAF dimensions. The findings further indicated that the influence of task complexity depended partly on task type, particularly for selected syntactic complexity measures.

Task type significantly influenced all three CAF dimensions, though its effects varied across measures. Decision-making tasks elicited greater syntactic complexity than narrative tasks, as reflected in MLT, CN/T, and CN/C, whereas no difference emerged for DC/T. These findings suggest that decision-making tasks primarily promoted phrasal elaboration rather than greater subordination. Within genre-based approaches (Biber et al., 2011; Norris & Ortega, 2009), this pattern aligns with the communicative demands of decision-making tasks, which require learners to evaluate alternatives and justify decisions, potentially leading to denser nominal constructions than in narrative writing. However, because the study examined linguistic products rather than cognitive or discourse processes, this explanation should be regarded as interpretive rather than causal.

The findings generally align with previous research demonstrating that writing task type shapes syntactic performance (Bao, 2025; Zhao et al., 2024, 2025). They also support studies reporting greater syntactic elaboration in reasoning- or decision-based writing (Allaw, 2019). At the same time, discrepancies with studies favoring narrative or descriptive writing (Garcia-Ponce & Tavakoli, 2022; Hsieh, 2024) likely reflect differences in learner characteristics, task design, and operationalizations of syntactic complexity.

The present study indicated that task type also significantly affected writing accuracy. Decision-making tasks produced higher error-free ratios and more accurate verb forms than narrative tasks. A possible explanation is that differences in discourse structure, linguistic demands, and communicative purpose account for the differences. Narrative writing requires sustained management of temporal reference, event sequencing, and referential cohesion, creating more opportunities for grammatical errors. This interpretation is consistent with previous studies reporting greater accuracy in decision-making tasks (Garcia-Ponce & Tavakoli, 2022) and more verb-related errors in narrative writing (Diez-Bedmar & Thewissen, 2025). However, the present data do not directly reveal the mechanisms underlying these differences.

In contrast, narrative tasks produced greater writing fluency than decision-making tasks. The more linear discourse structure of narratives may allow learners to sustain their production more easily than tasks that require evaluation and justification. Similar task-type differences in fluency have been reported in previous research (Garcia-Ponce & Tavakoli, 2022; Hsieh, 2024; Zhao et al., 2025).

Overall, these findings suggest that task type does not uniformly enhance writing performance but directs learners toward different linguistic resources. They therefore reinforce the central role of task design in shaping written production within TBLT.

Task complexity also exerted different effects across the three CAF dimensions. Increasing reasoning demands, which are associated with enhanced syntactic complexity and reduced accuracy, had no significant effect on fluency, suggesting that different aspects of writing respond differently to increased task demands.

Regarding syntactic complexity, learners produced longer T-units and greater phrasal elaboration (CN/T and CN/C) in the complex tasks, whereas subordination (DC/T) remained unchanged. This pattern broadly aligns with Robinson's Cognition Hypothesis, which posits that resource-directing increases in task complexity encourage more elaborate language production. Although the present study did not directly examine learners' cognitive processes, the findings indicate that more cognitively demanding tasks were associated with greater syntactic elaboration in the written product. The absence of an effect on DC/T further highlights the multidimensional nature of syntactic complexity and suggests that learners responded to greater task demands primarily through nominal elaboration rather than clause embedding (Biber et al., 2011; Norris & Ortega, 2009).

The present results align with studies reporting positive effects of task complexity on syntactic complexity (Allaw, 2019; Rahimi & Zhang, 2018) and with theoretical accounts that emphasize deeper semantic processing under increased cognitive load (Robinson, 2001, 2007). However, the literature also reports conflicting findings. Kuiken and Vedder (2007, 2008), Xu et al. (2023), and Tao and Wang (2025) reported limited, null, or even negative effects of task complexity on syntactic complexity. These discrepancies likely reflect differences in learner proficiency, task design, and the operationalization of syntactic complexity.

In contrast, increasing task complexity significantly reduced writing accuracy. Learners produced lower error-free ratios and fewer correct verb forms in the complex condition, a pattern more consistent with Skehan's Limited Attentional Capacity Model than with Robinson's Cognition Hypothesis. These findings suggest that higher reasoning demands may limit learners' ability to maintain grammatical accuracy while producing more syntactically elaborate language. Nevertheless, because the study analyzed written products rather than attentional processes, this interpretation remains theoretical rather than causal.

The observed decline in accuracy aligns with studies that report trade-offs as task complexity increases (Rahimi & Zhang, 2018, 2019; Rashtchi & Mohammad Yousefi, 2017; Zhan et al., 2021). At the same time, some studies found increased accuracy on complex tasks (Kuiken & Vedder, 2007, 2008; Robinson, 2005), while others reported no significant effects (Robinson, 2007; Xu et al., 2023). Overall, the mixed literature suggests that the relationship between task complexity and accuracy depends on multiple contextual and methodological factors.

Unlike syntactic complexity and accuracy, fluency was unaffected by task complexity. A possible explanation lies in the nature of written language production. Because writing allows learners to regulate planning, monitoring, and revision, increased cognitive demands may influence the linguistic characteristics of the completed text without substantially affecting production rate. This interpretation aligns with numerous studies reporting limited or non-significant effects on fluency (Rahimi & Zhang, 2018; Xu et al., 2023). However, previous research has also reported greater fluency in complex tasks (Zhan et al., 2021), greater fluency in simple tasks (Abrams, 2019), and fluency gains through task sequencing rather than complexity manipulation alone (Tabari & Miller, 2021), further underscoring the sensitivity of fluency outcomes to task design and measurement decisions.

Interaction analyses showed that the effects of task complexity varied by task type for selected syntactic complexity measures. Significant Task Type × Task Complexity interactions were observed for MLT, CN/T, and CN/C, whereas no interactions emerged for DC/T, accuracy, or fluency. These findings indicate that the impact of increasing task complexity was not uniform across writing tasks but varied by task type and the syntactic complexity dimension examined.

For MLT and CN/T, the increase associated with task complexity was greater in the decision-making task than in the narrative task, whereas the opposite pattern was observed for CN/C, with a larger complexity-related increase in the narrative task. These contrasting patterns reinforce the view that syntactic complexity is multidimensional and that different indices may respond differently to the same task manipulation. Accordingly, interaction effects should be interpreted at the level of individual CAF measures rather than as evidence of a general enhancement of syntactic complexity.

These findings are broadly consistent with Robinson's Cognition Hypothesis, which predicts that the effects of task complexity may vary with task characteristics. They also align with previous studies showing that task complexity does not affect all linguistic dimensions uniformly (Alexopoulou et al., 2017; Allaw, 2019). However, because the observed interaction effects were small to moderate and the study examined written products rather than cognitive processes, the results should not be interpreted as evidence for the attentional or processing mechanisms proposed by the Cognition Hypothesis. Instead, they demonstrate that task type moderated the effects of task complexity on selected measures of syntactic complexity under the conditions of the present study.



## 6. Conclusion

The present study examined the independent and interactive effects of task type and task complexity on EFL learners' written performance across multiple dimensions of CAF. The findings showed that task design influenced writing performance selectively rather than uniformly. Decision-making tasks elicited longer T-units, greater nominal elaboration, and higher accuracy, whereas narrative tasks produced greater written fluency. Increasing task complexity promoted phrasal syntactic complexity but reduced accuracy, with no significant effects on subordination or fluency. Significant Task Type × Task Complexity interactions for MLT, CN/T, and CN/C further indicated that the effects of task complexity varied by task type for selected aspects of syntactic complexity. These findings highlight the multidimensional nature of L2 writing. Different CAF dimensions responded differently to the same task manipulations, indicating that no single task condition optimized all aspects of performance. The results therefore reinforce the importance of evaluating writing using multiple CAF measures rather than relying on a single performance dimension.

The findings offer qualified support for both Robinson's Cognition Hypothesis (2001, 2003, 2005) and Skehan's Limited Attentional Capacity Model (1998, 2015). Although increased task complexity promoted syntactic elaboration, it was also associated with reduced accuracy, suggesting that cognitive engagement and attentional constraints jointly shape written performance. However, because the study examined linguistic products rather than cognitive processes, these interpretations should be viewed as explanatory rather than causal.

Pedagogically, the findings suggest that task design should be aligned with instructional objectives. Decision-making tasks may be appropriate when the goal is to promote syntactic elaboration and accuracy, whereas narrative tasks may better support fluent written production. Likewise, increasing task complexity can encourage more sophisticated language use but may require instructional support to minimize reductions in accuracy. Overall, the findings underscore the importance of balancing task type and cognitive demands when designing task-based writing activities.

Several limitations should be considered when interpreting the findings. First, the study relied exclusively on quantitative analyses of written products. Consequently, explanations involving attentional allocation, cognitive load, or planning processes remain inferential because these mechanisms were not measured directly. Second, only two task types (narrative and decision-making) and one operationalization of task complexity were examined. Thus, the findings should not be generalized to other task genres or complexity dimensions without further evidence. Third, although significant Task Type × Task Complexity interactions were identified, the observed interaction effects were generally small to moderate and should be interpreted cautiously when considering their practical implications. Fourth, the study was conducted in a single private language school and included only female EFL learners. Consequently, the findings may not be generalizable to learners in other institutional settings or to male or mixed-gender populations. Finally, task type was assigned at the level of intact classes rather than individual participants. Although all classes were taught by the same instructor using identical instructional conditions, the clustered design may have introduced some degree of non-independence among observations. Because the study included a limited number of classes, multilevel modeling was not feasible.

Therefore, future research employing a larger number of classes should use multilevel analyses to account explicitly for classroom-level effects. Additionally, more diverse learner populations and a wider range of writing tasks should be examined to determine the generalizability of the present findings. Additional studies should also investigate alternative operationalizations of task complexity and explore whether similar patterns emerge across other genres and language modalities. Incorporating process-oriented methodologies, such as keystroke logging, eye tracking, or stimulated recall, would provide direct evidence of the cognitive processes underlying task performance and strengthen theoretical interpretations of task effects.

## References

Abrams, Z. I. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task-complexity. *System, 81*, 110–121. https://doi.org/10.1016/j.system.2019.01.009

Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning, 67*(1), 180–208. https://doi.org/10.1111/lang.12232

Allaw, E. (2019). A learner corpus analysis: Effects of task complexity, task type, and L1 & L2 similarity on propositional and linguistic complexity. *International Review of Applied Linguistics in Language Teaching, 59*(4), 569–604. https://doi.org/10.1515/iral-2018-0294

Bao, G. (2025). Phrasal complexity in EFL learners' oral performance: The role of task type and proficiency, *64*(2), 869–903. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2024-0273

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35. https://doi.org/10.5054/tq.2011.244483

Bui, G., Skehan, P., & Wang, Z. (2018). Task condition effects on advanced-level foreign language performance. In P. Malovrh & A. Benati (Eds.), *The handbook of advanced proficiency in second language acquisition* (pp. 219–237). Wiley. https://doi.org/10.1002/9781119261650.ch12

Diez-Bedmar, M. B., & Thewissen, J. (2025). Exploring task type effects on L2 accuracy profiles: Narrative vs. email writing by B1 Spanish EFL learners. In S. De Cock, G. Gilquin, S. Granger, P. Jadoulle, & M. Paquot (Eds), *Book of abstracts: Register and task variation in learner corpus research* (pp. 19–20). Université catholique de Louvain.

Ellis, R. (2003a). Designing a task-based syllabus. *RELC Journal, 34*(1), 64–81. https://doi.org/10.1177/003368820303400105

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

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

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

Gilabert, R. (2005). *Task complexity and L2 narrative oral production* [Unpublished doctoral dissertation, University of Barcelona, Barcelona, Spain].

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and [+/– Here-and-Now]: Effects on L2 oral production. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Multilingual Matters.

Gilabert, R., Baron, J, & Llanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching*, *47*(3-4), 367–395. https://doi.org/10.1515/iral.2009.016

Gilanlioglu, I. (2022). Influence of planning time and task type on lexis in L2 oral performance. *Bartın University Journal of Faculty of Education*, *11*(3), 559–575. https://doi.org/10.14686/buefad.1026153

González-Torres, P. F., & Mayo, M. D. P. G. (2025). Grammatical transfer errors in EFL writing: Impact of Spanish influence, proficiency levels, and task types. *International Journal of Learning, Teaching and Educational Research*, *24*(5), 360–375. https://doi.org/10.26803/ijlter.24.5.19

Hsieh, C. N. (2024). The Role of task types and reading proficiency on young English as a foreign language learners' writing performances. *TESOL Quarterly*, *58*(2), 978–990. https://doi.org/10.1002/tesq.3286

Jiang, L., Abbuhl, R., & Kang, M. (2025). Effects of working memory and proficiency on lexical complexity in L2 writing: A case of two task types. *International Journal of Applied Linguistics*, *36*(1), 477–487. https://doi.org/10.1111/ijal.12791

Jin, C., & Yan, J. (2025). The effects of task complexity and task sequencing on L2 performance: A systematic review. *The Language Learning Journal*, *53*(1), 114–141. https://doi.org/10.1080/09571736.2024.2345911

Kim, N. (2023). The effects of different task sequences on novice L2 learners' oral performance in the classroom. *Language Teaching Research*, *27*(2), 415–440. https://doi.org/10.1177/1362168820937548

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching, 45*(3), 261–284. https://doi.org/10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1)*,* 48–60. https://doi.org/10.1016/j.jslw.2007.08.003

Lee, J. (2020). Task closure and task complexity effects on L2 written performance. *Journal of Second Language Writing, 50*, Article 100777. https://doi.org/10.1016/j.jslw.2020.100777

Liu, J., Shi, Z., & Li, W. (2025). Understanding genre differences in cohesive use of EFL Learners: Insights from systemic functional linguistics. *System*, *132*, Article 103692. https://doi.org/10.1016/j.system.2025.103692

Malicka, A. (2020). The role of task sequencing in fluency, accuracy, and complexity: Investigating the SSARC model of pedagogic task sequencing. *Language Teaching Research, 24*(5), 642–665. https://doi.org/10.1177/1362168818813668

McDonough, K., & Fuentes, C. G. (2015). The effect of writing task and task conditions on Colombian EFL learners' language use. *TESL Canada Journal, 32*(2), 67–79. https://doi.org/10.18806/tesl.v32i2.1208

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578. https://doi.org/10.1093/applin/amp043

Pang, F., & Skehan, P. (2021). Performance profiles on second language speaking tasks. *The Modern Language Journal, 105*(1), 371–390. https://doi.org/10.1111/modl.12699

Polio, C. (1997). Measures of linguistic accuracy in second-language writing research. *Language Learning, 47*(1), 101–143. https://doi.org/10.1111/0023-8333.31997003

Qiu, X., & Cheng, H. (2022). The effects of task types on L2 oral production and learner engagement. *International Review of Applied Linguistics in Language Teaching*, *60*(4), 1063–1088. https://doi.org/10.1515/iral-2020-0128

Rahimi, M., & Zhang, L. J. (2018). Effects of task complexity and planning conditions on L2 argumentative writing production. *Discourse Processes, 55*(8), 726–742. https://doi.org/10.1080/0163853X.2017.1336042

Rahimi, M., & Zhang, L. J. (2019). Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Reading and Writing*, *32*(3), 761–786. https://doi.org/10.1007/s11145-018-9887-9

Rashtchi, M., & Mohammad Yousefi, L. (2017). Reading input flooding versus listening input flooding: Can they boost speaking skill? *Journal of Language and Cultural Education (JoLaCE), 5(*1), 39–58. https://doi.org/10.1515/jolace-2017-0003

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, *38*(4), 703–737. https://doi.org/10.1017/S0272263115000339

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Studies in Second Language Acquisition, 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 in Language Teaching, 43*(1), 1–32. https://doi.org/10.1515/iral.2005.43.1.1

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL, 45,* 193–213. https://doi.org/10.1515/IRAL.2007.009

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 the development of TBLT* (Vol. 8, pp. 87–121). John Benjamins.

Ryu, J., & Lee, B. (2018). Effects of task complexity and planning on L2 speaking performances and focus on form. *Korean Journal of Applied Linguistics, 34*(4), 255–286. http://dx.doi.org/10.17154/kjal.2018.00.34.4.255

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38–62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2015). Limited attentional capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–155). John Benjamins.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics*, *36*, 34–49.

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

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge University Press.

Skehan, P., & Foster, P. (2005). Strategic and online planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task-performance in a second language* (pp. 193–218). John Benjamins.

Tabari, M. A., & Lee, J. (2025). 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

Tabari, M. A., & Miller, M. (2021). Unraveling the effects of task sequencing on the syntactic complexity, accuracy, lexical complexity, and fluency of L2 written production. *Canadian Journal of Applied Linguistics*, *24*(2), 1–29. https://doi.org/10.37213/cjal.2021.31306

Tao, N., & Wang, Y. (2025). Effects of prior knowledge and reasoning demands on Chinese EFL writing performance. *Language Teaching Research*, *29*(3), 1264–1285. https://doi.org/10.1177/13621688221089879

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, *67*(2), 394–430. https://doi.org/10.1111/lang.12228

Xu, T. S., Zhang, L. J., & Gaffney, J. S. (2023). A multidimensional approach to assessing the effects of task complexity on L2 students' argumentative writing. *Assessing Writing*, *55*, 1–16. https://doi.org/10.1016/j.asw.2022.100690

Yoon, H. J. (2021). Challenging the connection between task perceptions and language use in L2 writing: Genre, cognitive task complexity, and linguistic complexity. *Journal of Second Language Writing*, *54*, Article 100857. https://doi.org/10.1016/j.jslw.2021.100857

Zhan, J., Sun, Q., & Zhang, L. J. (2021). Effects of manipulating writing task complexity on learners' performance in completing vocabulary and syntactic tasks. *Language Teaching Research*, 1–22. https://doi.org/10.1177/13621688211024360

Zhao, Z., Nimehchisalem, V., & Chan, M. Y. (2024). Independent writing tasks vs. integrated writing tasks: The cognitive demands and their impacts on linguistic complexity in EFL writing. *Theory & Practice in Language Studies (TPLS)*, *14*(8), 2606–2617. https://doi.org/10.17507/tpls.1408.33

Zhao, Z., Nimehchisalem, V., & Chan, M. Y. (2025). The effects of task types and working memory capacity on the writing production of Chinese EFL learners. *Journal of Language Teaching & Research*, *16*(2). 515–525. https://doi.org/10.17507/jltr.1602.17