Does the Relationship between Executive Functions and L2 Writing Depend on Language Proficiency?

Lisa Haake, Leibniz University Hannover | Germany

Abstract: Executive functions are attributed a central role in maintaining fluency during L2 text composition, allowing writers to orchestrate the various linguistic and cognitive processes and resources involved in writing. The study examined (1) whether language proficiency moderates the relationship between executive functions and writing fluency in L2 writing and (2) whether the effects indirectly affect text quality. Sixty university students composed two texts in English as their L2, an argumentation and a description, three executive function tasks assessing inhibition, shifting, and updating skills, a language proficiency test, and a copy task. Keystroke logging protocols were recorded with Inputlog and analyzed for writing fluency. Text quality was assessed with a holistic benchmark procedure and comparative judgments. The results revealed language-dependent and genre-specific effects of updating and shifting but not inhibition skills on writing fluency. Path models indicated that the interactions between executive functions and language proficiency indirectly affect text quality through process-related writing measures. The findings suggest a complex relation between executive functions and writing performance that depends on language proficiency and varies with task demands.

Keywords: L2 writing; executive functions; language proficiency; writing process; keystroke logging

journal of
WRITING RESEARCH

Haake, L. (2025). Does the relationship between executive functions and L2 writing depend on language proficiency? *Journal of Writing Research*, *17*(1), 115-153. DOI: https://doi.org/10.17239/jowr-2025.17.01.05

Lisa Haake, Leibniz University Hannover, Institute of Psychology, Im Moore 11, 30167 Hannover | Germany – haake@psychologie.uni-hannover.de – Orcid: 0000-0001-8195-4022

1. Introduction

Writing requires the skillful coordination of various linguistic and cognitive processes and resources. Efficient cognitive processing and well-developed language skills are posited pivotal roles for orchestrating and implementing the different writing sub-processes in second language (L2) text composition (Kormos, 2012, 2023). Executive functions (EFs) have received increased attention in researching the inter-individual differences accounting for differences in L2 writing performance (Manchón et al., 2023; Mavrou, 2020; Vallejos, 2020; Rйvйsz et al., 2023). EFs refer to a set of cognitive processes that control, coordinate, and regulate the performance of complex cognitive tasks (Diamond, 2013; Friedman & Miyake, 2017; Miyake et al., 2000). Regarding writing, they are assumed to enhance the intercommunication between motor and sensory skills, linguistic resources, and higher-order cognitive tasks, such as reasoning, problem-solving, and planning, as they engage in the writing process. More specifically, EF skills are assumed to facilitate lexical retrieval processes by inhibiting lexical competitors. They allow for skillfully shifting between writing sub-processes and task demands. They support maintaining and updating linguistic formulations in WM (Olive, 2021, 2022).

In first language (L1) writing, the critical role of EFs is theoretically well-justified (for a published handbook on the relations between EFs and L1 writing, see Limpo & Olive, 2021) and supported by empirical research on beginning (Cordeiro et al., 2020; Costa et al., 2020; Drijbooms et al., 2015, 2017) and experienced writers (Larigauderie et al., 2020; Olive, 2004). Due to non-automated and incomplete linguistic knowledge, L2 writing is supposedly more effortful and requires more attentional resources than L1 writing (Kormos, 2012, 2023). While limited language skills may thus increase the significance of efficient cognitive processes, empirical research on the relationship between EFs, language proficiency, and L2 writing is scarce, and the role of language proficiency remains largely unknown. The existing body of research on the role of EFs in L2 writing yields rather inconsistent findings (for a review, see Li, 2023). The difficulty in comparing findings between samples is compounded by the fact that writers with different language proficiency levels compose texts of different genres and under different task settings. The present study aimed (1) to investigate the role of language proficiency in the relationship between EFs and L2 writing fluency and (2) to examine indirect effects on text quality. By extending the research to process-related measures of writing performance in two task genres, i.e., argumentative and descriptive writing, the findings seek to contribute to a more comprehensive understanding of what accounts for students' differential success in L2 text composition.

2. Theoretical background and literature review

Writing is a recursive process involving planning, translating, transcribing, and revising (Hayes, 2012). Writing fluency can be defined as the ability to convert non-verbal information into verbal form quickly, efficiently, and without significant hesitation (Dux Speltz & Chukharev-Hudilainen, 2021; Linnemann et al., 2022). It is a central construct in L2 writing, characterizing writing performance, predicting text quality, and capturing language proficiency (Abdel Latif, 2013; Van Waes & Leijten, 2015; Skehan, 2009). In writing process research, fluent writing processes have been characterized by high production rates, few and short pausing times, few revisions, and long bursts (Van Waes & Leijten, 2015). To maintain writing fluency, writers must operate the various components of text composition simultaneously (Olive, 2014). Low-level writing processes, such as translating and transcribing, must be sufficiently automated to free up cognitive resources for more complex, high-level writing activities, such as planning or revising, given the limited capacity of the cognitive system. When translation and transcription processes demand excessive attention, in return, writers may struggle to manage the various attention-binding tasks involved in writing, increasing the risk of cognitive overload. This can slow down writing, lead to interruptions in typing, or result in errors (Chenoweth & Hayes, 2001). The availability of linguistic and cognitive resources is assumed to influence the coordination and execution of these writing sub-processes, ultimately affecting the quality of the final text (Kormos, 2012, 2023). The following sections will explore the relationships between these variables in greater depth.

2.1 The influence of linguistic resources on writing

Among experienced writers composing a text in a well-known L1, low-level processes are usually automated, allowing them to devote attention to high-level activities while they execute the low-level processes (Olive, 2014). If linguistic knowledge, such as vocabulary, syntactic, or pragmatic knowledge, is incomplete and insufficiently automated – as usually is for L2 writers – also low-level processes become effortful. Depending on the language proficiency level, L2 writers compose texts with relative fluency: Greater availability of automatic access to linguistic knowledge in an L2 frees up cognitive resources for higher-order writing activities, which allows for fluent writing processes and, in turn, positively affects the language, content, and structure of a text (Kormos, 2023). The ability of L2 writers to compose texts fluently is contingent upon the level of complexity inherent to the writing task. While proficient L2 writers may be capable of producing simple texts with a certain degree of fluency, they may require periodic pauses for reflection when confronted with more intricate texts (Vandermeulen et al., 2024).

Generally, well-developed language skills cohere with high text quality and fluent writing skills (e.g., Haake et al., 2024). Writers with high language proficiency tend to produce language at higher rates than low-proficient writers, indicating increased automaticity with which writers implement the writing processes (Chenoweth & Hayes, 2001). Well-developed language skills further allow for the planning of longer linguistic sequences reflected in fewer interruptions of the writing flow. Writers with low language proficiency, in contrast, frequently produce pauses at low thresholds (e.g., 200 ms) and pauses between words, indicating extensive planning of small linguistic units. As language skills evolve, pauses shift to sentence boundaries and higher thresholds (e.g., 2000 ms), suggesting that more automated access to linguistic knowledge allows for more global text planning (Chukharev-Hudilainen et al., 2019; Révész et al., 2022; Barkaoui, 2019). Extensive revising may result from inefficient transcription processes or attempts to improve or correct the employed language, content, or structure (Galbraith & Baaijen, 2019). It has been observed that both low- and high-proficient writers may engage in extensive revisions. While writers at low proficiency levels are typically concerned with revising linguistic aspects of the text, writers at advanced language levels tend to increasingly revise structural features of the text (Barkaoui, 2016; Lindgren et al., 2008).

2.2 The influence of cognitive resources on writing

In addition to having access to sufficiently automated linguistic knowledge, writers must have sufficient cognitive resources available to compose a text fluently (Olive, 2014). EFs are assumed to play a pivotal role during all stages of the writing process, allowing a writer to skillfully orchestrate and implement the various attention-binding sub-tasks involved in writing (Kellogg, 1996; Kellogg et al., 2013; Kormos, 2022, 2023; Olive, 2021, 2022). According to Miyake et al. (2000), EFs include three distinguishable but correlated functions, i.e., inhibition, shifting, and updating. They are understood as cognitive sub-components of one supervisory attentional mechanism that interact, resulting in a trade-off between resources devoted to each function. The EF subcomponents are attributed different functions in implementing and orchestrating the distinct writing sub-processes, collectively supporting writers to achieve high-quality writing performance (Olive, 2021).

Inhibition entails the skills of inhibiting prepotent responses, attending to task-relevant while suppressing task-irrelevant information, and completing a task despite distractors. In terms of writing, inhibition skills are argued to be particularly crucial for retrieving knowledge from long-term memory. As dominant or prepotent responses are more prone to be activated in memory, writers must inhibit simpler words and formulations frequently used in speech, for instance, to employ more complex language (Kellogg et al., 2013). During L2 production, efficient inhibition skills are thought to facilitate the access and final utterance of correct lexical competitors of an intended language (i.e., the less dominant L2) and, in return, the suppression of competitors of the unwanted language (i.e., the more dominant L1; Kormos, 2012). Inhibition skills are assumed to be particularly important for low-proficient L2 learners when access to the relevant linguistic knowledge is not sufficiently automated and L1 influence is greater (Granena, 2023; Li, 2023).

Shifting refers to the skill of flexibly switching between mental sets. Shifting is thought to promote switching between the various attention-binding tasks involved in text composition, allowing a writer to adapt efficiently and quickly to new and varying task demands, thoughts, and goals throughout the writing process (Olive, 2021). Shifting processes support, for instance, shifting from process to content planning, between knowledge when ideas are translated into language, from sentence generation to revising, or from global to local revising processes (Drijbooms et al., 2015, 2017). L2 writers have been shown to fall back on (meta-) linguistic knowledge of their L1 to solve emerging problems in an L2 (e.g., Van Weijen et al., 2009). Since switching between languages results in a cost (Gade et al., 2021), efficient shifting processes might further facilitate and support knowledge transfer from one language to another.

Updating describes the skills of temporarily storing, manipulating, renewing, and monitoring information in working memory. Updating is argued to allow writers, among others, to keep in mind the writing task for content generation (planning), remember ideas sought to be converted into language (translating), memorize a linguistic string until it has been written down (transcribing), or compare the initial goals of the writing task to the written output (reviewing; Olive, 2021). Efficient updating processes allow L2 writers to temporarily store retrieved knowledge from long-term memory to process, update, and integrate it with the transient information resulting from the execution of the distinct writing processes. Producing complex language in an L2 requires maintaining and updating linguistic formulations in WM. Thus, efficient updating processes may facilitate employing complex language (Mavrou, 2020).

2.3 Empirical findings on the relationship between executive functions, language proficiency, and writing

Studies investigating the process-related aspects of writing have reported inconsistent findings. A study conducted by Kim and colleagues (2021) lends support to the involvement of EFs in the writing process. To investigate the influence of cognitive and linguistic resources on writing processes and products, the researchers collected data from L2 learners of English from diverse language backgrounds in the US, who composed a persuasive writing task. Those with superior inhibition skills[1] paused significantly less and shorter than their counterparts. This, in turn, was found to be associated with a higher level of writing quality. The findings indicate that the capacity to attend to task-relevant information while suppressing task-irrelevant details enables writers at advanced language proficiency levels to compose a text with reduced cognitive effort. However, neither the direct effect of inhibition skills on writing quality nor the indirect effect through pausing reached statistical significance. In addition, inhibition skills were not associated with other measures of writing fluency, such as burst length or production rates. Révész et al. (2017) provide further evidence for a positive effect of cognitive resources on L2 writing fluency. The researchers investigated how individual differences in EF skills among Mandarin learners of English affect writing processes and text quality in argumentative writing. Although EFs did not correlate with text quality, higher updating and shifting skills significantly correlated with shorter pauses between paragraphs in argumentative writing. Additionally, updating was positively related to production rates and revision features, indicating that efficient EFs facilitate more fluent compositions of argumentative texts.

Vallejos (2020) examined the pausing behavior of emergent English-Spanish bilinguals composing argumentative texts in English as their L1 and Spanish as their L2. While superior shifting skills were not associated with either pause frequency or length in either language, the researcher observed that better updating skills were related to longer and more frequent pauses between sentences in L2 writing, but not in L1 writing. The findings indicate that language proficiency might be crucial for the relationship between EF skills and writing behavior. Contrary to expectations, however, the findings suggest that better updating skills are associated with reduced, rather than enhanced, writing fluency. Similarly, Torres (2023) found that updating skills[2] influence online pausing and revision behaviors of Spanish-English heritage bilinguals in argumentative writing. Using think-aloud protocols, the researcher found that better updating skills were associated with longer within-word pauses, rather than shorter ones. He speculated that efficient updating skills allow writers to devote more attention to encoding issues, which was reflected in the longer pauses. This was supported by the collected think-aloud data. In her review on the effects of individual differences on writing, Granena (2023) suggests that writers might use long pauses for more accurate linguistic encoding, which could result in higher-quality texts.

These findings are consistent with those of Révész et al. (2023), who investigated the effects of EF skills, among other cognitive variables, on pausing and revision behaviors across different stages of writing. The researchers collected data from advanced Chinese L2 learners of English composing an argumentative essay. While inhibition skills were not found to influence writing behavior, superior shifting skills were associated with a greater frequency of within-word pauses (with a threshold of 2000 ms) during the middle stages of writing. The results indicate that writers with better shifting abilities utilize their enhanced skills for frequent switching between higher- and lower-order writing processes. Superior updating skills were further associated with shorter pauses at sentence boundaries at the later stages of writing, in which writers tend to predominantly monitor their writing. This suggests that better updating skills facilitate more efficient coordination and updating of monitoring operations. In line with Torres' findings, EF skills did not affect the revision behavior, which the researchers attribute to the possibly low cognitive demands of the implemented writing task and the high L2 and writing proficiency of the sample.

In his systematic review on the relationship between working memory and L2 writing, Li (2023) identifies several third variables that may act as moderators for the effects of cognitive resources on writing performance, including language proficiency and task demands. In their seminal work, Kormos and Sáfár (2008) demonstrated that the influence of cognitive resources on L2 skills varies at the beginning and pre-intermediate stages of language learning. The researchers collected data from Hungarian adolescents, who were learning English as an L2. They completed an L2 language proficiency test that evaluated writing proficiency, along with other L2 performance areas, including reading, listening, and speaking. A positive correlation was observed between phonological short-term memory and overall L2 proficiency, the use of English, and writing proficiency for learners at a pre-intermediate language level, but not at a beginning language level. Additionally, the researchers investigated the updating skills of the beginning language learners, which did not significantly contribute to performance differences in writing quality. These findings indicate that individual differences in cognitive resources do not explain performance differences in L2 writing at the initial stages of language learning. These differences only emerge once a certain language level has been reached.

Further evidence for a language-dependent shift of EF effects between performance aspects of writing is provided by a study conducted by Vasylets and Marнn (2021). The researchers collected data from Spanish university students, who wrote a narrative text in English as their L2. Although updating skills had no significant effect on overall text quality, enhanced updating skills were related to higher accuracy for less proficient writers and to more sophisticated language for more proficient writers. It appeared that higher updating skills compensated for limited language proficiency in some language domains, i.e., in the production of error-free language. In other domains, namely the production of complex language, only students with superior language abilities benefited from enhanced updating capacities. In a recent study, Manchón et al. (2023) examined the interactive effects between language proficiency and updating on the L2 writing performance of Spanish undergraduate students learning English. While language proficiency had a significant effect on complexity, accuracy, and product-related fluency measures, neither updating nor interactive effects between language proficiency and updating have been observed to have a significant effect. These results differed from those reported by Vasylets and Marнn, which the researchers attributed to their task settings with relatively long time limits, in which writers could solely rely on their linguistic and literacy knowledge without a detectable variance in updating skills.

A study by Leong and colleagues (2019) provides further evidence for the moderating effect of task demands on the relationship between EF skills and text characteristics. The researchers reported that updating skills were predictive of text quality in explanatory and argumentative writing, but not in narrative writing, for young L2 Chinese learners with Urdu or Hindi as an L1. The researchers explained the findings with the varying degrees of complexity that text genres place on writers. The cognitive demands of composing an explanation and argumentation are greater as those of narrative writing, as writers must store and update multiple pieces of information to compose texts that are meaningful and coherent. In contrast, narrative writing typically adheres to a chronological order, which facilitates coherence building.

The findings are consistent with those of Zalbidea (2017), who examined the effects of task complexity, manipulated via reasoning demands in argumentative writing. The researcher found that updating was associated with accuracy in complex task conditions but not in simple task conditions for English learners of Spanish. Although a study by Cho

(2018) did not confirm the interactive effects, the findings indicate that task complexity is a crucial factor in the relationship between EF skills and writing behavior. It seemingly must be sufficiently challenging for cognitive abilities to affect writing performance.

The studies conducted so far have made promising attempts at elucidating the potential involvement of sub-components of EFs in the writing process. Several conclusions can be drawn from the literature reviewed: Firstly, evidence suggests that EFs exert no direct influence on global measures of text quality. It seems probable that such integrated quality measures are not sensitive enough to detect cognitive effects on writing performance (see also Li, 2023). Secondly, the literature indicates that EFs are critical for writing as a function of language proficiency. The availability of linguistic resources has been shown to influence the extent to which writers benefit from enhanced cognitive resources. It seems therefore imperative to assess language proficiency levels to investigate the effects of EF skills on writing performance. Thirdly, several studies indicate that task complexity has a significant influence on the extent to which cognitive resources influence writing performance. If the task settings are not sufficiently challenging, writers appear to be able to solve the tasks equally well, such that variability between cognitive resources becomes undetectable. Lastly, given the assumption that different sub-components of EFs are assumed to interact, resulting in a trade-off between resources devoted to each process, it seems necessary to examine not only a single sub-component of EF, but all three sub-functions, to reliably unveil the effects of EFs on writing performance under the given requirements. It is important to mention that methodological limitations in the investigation of the influence of attentional resources on L2 writing may have contributed to the inconclusive findings. Interested readers are directed to Li (2023) and Willoughby and Hudson (2021).

3. The present study

Motivated by the identified research gaps, the present study investigated how sub-components of EFs relate to fluent writing processes in argumentative and descriptive text composition, aiming to (1) determine the role of language proficiency in the relationship between EFs and L2 writing fluency and (2) examine indirect effects on text quality. Generally, writers with more efficient EF skills are expected to better cope with the linguistic and structural demands of the writing tasks than writers with less efficient EF skills. Efficient EF skills are thus expected to positively affect writing on a process level, which, in turn, affects overall writing quality. The direction and strength of the effects are further anticipated to vary with language proficiency levels. Depending on available linguistic resources, writers face distinct challenges during L2 text composition, which EF skills are assumed to compensate for. Due to incomplete and insufficiently automated access to linguistic knowledge, the formulation process may impose particular difficulties at low proficiency levels. At high proficiency levels, the formulation process is less attention-demanding, which allows writers to direct cognitive resources to cognitively higher-order writing processes. Thus, writers at advanced language proficiency levels might benefit from better EF skills during higher-order cognitive processes. Lastly, the different requirements of the writing tasks are expected to produce distinct effects of EFs on the writing process as a function of language proficiency. While argumentative writing requires deep processing and careful deliberation of arguments and allows for more complex vocabulary and sentence structures, descriptive writing calls for more simple and familiar forms of writing. Thus, both task genres present writers at different language proficiency levels with distinct challenges, which is anticipated to be reflected in the interactive effects between EF skills and language proficiency.

4. Method

4.1 Participants

The study reports data from 58 out of 60 students of different subjects at the Leibniz University Hannover (30 female, 28 male, M_{age} = 24.76 years, age range = 19 − 35, *SD* = 2.90). Two participants were excluded from the analyses, one due to color blindness and one due to exceedingly slow and low performance across all EF tasks. All participants speak German as an L1, with English being considered an L2. For all participants, German is an L1 acquired before age three, and English is an L2 initially learned mainly at school. Eleven participants (18.97 %) reported to have an additional L1. They all received their entire education in Germany with no or minor contact with English before school enrollment. All participants reported having normal or corrected to normal vision and confirmed not having other predispositions that could have affected performance in the experiment.

4.2 Writing tasks

Two standardized, independent writing tasks were employed, in which the participants were invited to write a descriptive and argumentative text in English as their L2. In the argumentative task condition, the students were asked to explain their opinion on a controversial topic, i.e., should short-haul flights be banned within the European Union? In the descriptive task condition, the students were requested to describe a visible and familiar object, i.e., describe the flat they currently live in (Appendix A). An argumentative and a descriptive task design were chosen since both are central writing genres in higher education. Both task types were designed to address different writing and language skills: While the descriptive task calls for concisely providing information about what is being described, allowing the reader to form an accurate picture, and may rely on simple vocabulary and sentence structure, argumentative writing requires deep processing and careful deliberation of arguments and allows for employing more complex language (Grabowski & Mathiebe, 2024). Both were assumed, as such, to be complex enough to reflect inter-individual differences in writing performance. The argumentative prompt reduced interference from topic familiarity by allowing expert and novice content knowledge (e.g., comparing the impact of CO2 emissions of driving and flying upon the environment versus evaluating the benefits of short trip holidays). The descriptive prompt largely excluded prior content knowledge influencing writing performance by relying on existing, familiar content knowledge. Implementing a description of a personal object and a discussion of a controversial topic affecting one's personal life intended to stimulate task processing (Weigle, 2002) and motivation to use one's resources (Becker-Mrotzek, 2022).

The participants were instructed to read the respective assignment and mentally prepare the content of their texts for one minute. After, participants had fourteen minutes to complete the argumentative text and seven minutes to complete the descriptive text. They received a notification of the time one minute before the end. For either task condition, a text length of 150 to 250 words was indicated as a guideline. A relatively short time limit was set to increase task difficulty and, in turn, demands on EFs (Manchón et al., 2023). A lower time limit was chosen for the description to account for differences in task complexity. The specific time limits and text lengths were based on the average writing time of a previous study with a similar sample and the same task (Haake, 2024) and pilots.

4.3 Measures

Writing fluency

Keystroke logging data was recorded with Inputlog 8 (Leijten & Van Waes, 2013). The keystroke logging protocols were generated via the integrated general analysis of Inputlog and subsequently loaded into R (R Core Team, 2021) for pre-processing and analyses. The log-files were individually trimmed to remove noise at the beginning and end of each writing process and time-filtered using a bottom threshold of 30 ms for the inter-key interval to exclude continuous key pressing and unintentional double strokes, following Van Waes et al. (2021). The cleaned log-files were used for a fine-grained analysis of various fluency characteristics.

Studies that use keystroke logging data to investigate writing fluency have frequently distinguished between three different groups of features: (1) features related to productivity, i.e., production rates defined as the characters or words produced per minute (Chenoweth & Hayes, 2001; Révész et al., 2022, Spelman Miller et al., 2008; Van Waes & Leijten, 2015), (2) features related to inter-key intervals (IKIs), i.e., the time that elapses between two subsequent keystrokes (Barkaoui, 2019; Chukharev-Hudilainen et al., 2019; Van Waes & Leijten, 2015; Révész et al., 2019; Révész et al., 2022), and (3) features related to revisions, i.e., deletions from the text and insertions to the text away from the leading edge (Barkaoui, 2016; Lindgren et al., 2008; Révész et al., 2017; Révész et al., 2019; Spelman Miller et al., 2008). Based on the literature, a set of complementary keystroke features was selected that allows for a nuanced assessment of writing fluency for writers with different language proficiency levels (Table 1):

(1) Production rates were analyzed at a product and process level. Production rates at a product level compare the final text length to the overall writing time. Process-based research showed that production rates at the process level (i.e., the total number of characters produced during the process, including deleted characters) provide valuable information about writing fluency as they also factor in revised characters (Van Waes & Leijten, 2015).

(2) Inter-key intervals were analyzed at lexical boundaries, i.e., between words and sentences, and related to the overall produced words and sentences, respectively. No pause threshold was set as a discussion has recently emerged about whether pauses with fixed thresholds reflect the inter-individual differences between writers appropriately (Hall et al., 2022). The inter-key intervals between words and sentences were summed and divided by the total number of words and sentences written in the process to determine the average inter-key interval between two words and sentences, respectively.

(3) Revising processes were assessed using the revision frequency and the burst length (i.e., revision burst or r-burst). A product-process comparison was included as a measure of writing efficiency (Grabowski, 2008). Therefore, the total number of revisions (i.e., consecutive backspaces and insertion events away from the leading edge) was related to the minutes of processing time. Further, the total number of characters produced in the process was related to the total number of revisions to determine the average number of characters produced between two revisions. For a product-process comparison, the total number of characters in the final text was related to the total number of characters produced during the process, where higher values indicate little and lower values indicate more revised text.

Table 1 lists all measures and their descriptions. A detailed description of the feature extraction procedure is provided in Appendix A.

Table 1: Writing process measures with description

Variable	Description
Characters per minute (product)	Total number of characters in the final product per minute of total writing time
Characters per minute (process)	Total number of characters produced in the process (including deleted characters) per minute of total writing time
Inter-key interval between words	Total time between words per total number of words produced in the process

Variable	Description
Inter-key interval between sentences	Total time between sentences per total number of sentences produced in the process
Revisions per minute	Total number of revisions (deletions and insertions away from the leading edge) per minute of total writing time
Characters per burst	Average number of characters produced between two revisions
Product-process ratio	Characters in the final product per characters produced in the process

Text quality

Three raters assessed the text quality holistically, using a benchmark procedure and comparative judgments to rank the texts from poor to good text quality. Similar procedures have previously revealed high inter- and intra-rater reliabilities in L1 writing (Bouwer et al., 2023) and L2 writing (Siekmann et al., 2022; Tillema, 2012; Van Weijen, 2008). All raters were native German speakers and highly proficient in English. They all work as teachers in academic settings, use English daily, and are familiar with the English language and its writing conventions. Benchmark texts were selected from previous data collections with the same prompt and a similar sample regarding age, education, and language background after inspecting the current set of texts. The benchmark texts represented the average quality for good, intermediate, and poor performance with an explanation. Orthographic and punctuation errors were corrected to avoid biased judgments (Van Weijen, 2008).

To preserve the characteristics of the texts, inappropriate vocabulary and grammar were not corrected. Three rating rubrics were established: organization (coherence and structure), content (argumentation: quality and validity of arguments, their logical introduction and relevance, references and/or rationales; description: descriptiveness, figurativeness, details, examples, and their relevance), and language (communicative purpose, effectiveness, verbal skill, and rhetorical power). The raters were advised to consider and give the same weight to all three rubrics. The rating procedure was conducted as follows: Firstly, the raters assigned the texts to one of the three performance categories, such that each category comprised an equal number of texts. After repeating the procedure within each category, the raters sorted the texts from poor to good performance with comparative judgments of adjacent texts. Lastly, the raters reviewed and confirmed the order or undertook the necessary changes. Each text received points for its list position, and quality scores were obtained using the average of all raters. In a revision phase, 10 % of the texts with the most disagreement between the raters were reevaluated. The intraclass correlation coefficient (ICC) indicated excellent agreement in the argumentation (ICC = .96, 95 % CI [.94,.97]) and description (ICC = .92, 95 % CI [.87,.95]).

Executive Functions

EFs were assessed with a Simon task (Simon & Wolf, 1963) for inhibition, a Color-Shape task (Rubin & Meiran, 2005) for shifting, and an Operation Span task (Unsworth et al., 2005) for updating skills, all implemented in E-Prime (Psychology Software Tools, 2016). The tasks were chosen since they are relatively content- and language-free to reduce the impact of confounding variables and ensure a largely unbiased assessment of EF skills. All three tasks are well-established paradigms to assess EFs (Clair-Thompson & Wen, 2021; Karr et al., 2018; Nyongesa et al., 2019). A detailed description of each task is provided in Appendix B.

The Simon task, a paradigm based on reaction times, requires the inhibition of irrelevant spatial information pertaining to a presented stimulus and the direction of attention to its color. Participants were required to respond to the color of a stimulus (i.e., green or red) with their right or left hand and ignore its relative location on the screen (i.e., left or right). The location and color are assumed to cause interference due to a fast but decaying response activation of the irrelevant spatial position and a slow activation of the relevant stimulus feature (Hommel, 1993). Solving the task requires the deliberate and controlled suppression of a prepotent response. Hence, performance is typically associated with the inhibition component of EFs, such that faster and correct responses relate to enhanced inhibition skills. The task has been demonstrated to have good internal consistency and test-retest reliability (Clair-Thompson & Wen, 2021; Lu & Proctor, 1995).

The Color-Shape task is a reaction time-based task-switching paradigm in which participants must respond to either the color or the shape of a presented stimulus depending on a cue that precedes it. To process the two competing tasks, information from the previous trial must be inhibited, and the new task set must be prepared (Allport & Wylie, 2000). As the task requires switching flexibly between the varying task demands, performance is attributed to the shifting component of EFs, such that superior shifting skills are assumed to be reflected in faster response times and higher accuracy. Joint reaction times and accuracy scores have good test-retest reliability (ICC = .68; Sicard et al., 2022).

During the Operation span task, participants must maintain sets of unrelated letters with varying lengths in memory while solving a series of mathematical equations. The math operations serve as a secondary task, preventing the rehearsal of the to-be-remembered letters. Solving the task requires temporarily storing and updating information, a skill attributed to the updating component of EFs (Clair-Thompson & Wen, 2021; Wilhelm et al., 2013), such that a higher number of correctly recalled letters relates to superior updating skills. The task has good internal consistency ($\alpha = .78$) and test-retest reliability ($r = .83$; Unsworth et al., 2005).

Language proficiency

Language proficiency was assessed with a standardized cloze test (c-test), administered and implemented by the Leibniz Language Center of the Leibniz University Hannover. The test consists of five authentic texts in which every second word misses half of the letters (for a detailed description, see www.c-test.de). The texts have varying levels of difficulty and are calibrated and benchmarked, following the standards of the Association of Language Testers in Europe. They are selected randomly from a database and participants must fill in the gaps semantically, orthographically, and grammatically correctly. The test returns the percentage of correctly filled-in gaps as a continuously varying measure with cut scores for achieved language competence levels after the Common European Framework of References for Languages (CEFR; Council of Europe, 2001). A1 and A2 ($\leq$ 47 %) relate to a basic, B1 and B2 (48-84 %) to an intermediate, and C1 and C2 ($\geq$ 85 %) to a proficient language level. The test provides an objective, reliable, and valid way of measuring general language competence that usually correlates highly with more elaborate tests (for a review, see Eckes & Grotjahn, 2006).

Transcription skills

Participants performed the standardized English Inputlog copy task (for a detailed description, see Van Waes et al., 2019) to assess their transcription skills. The task targets typing and motor skills objectively, reliably, and validly (Van Waes et al., 2021). Two measures were computed and used as covariates: The median inter-key interval served as an indicator for general typing speed, and the percentage of correctly copied characters for typing accuracy.

Procedure

Each participant completed two writing assignments (i.e., a description and an argumentation), three EF tasks, an English placement test, a copy task, and a biographical background questionnaire. Two participants were tested simultaneously. The participants sat on opposite sides of the room, facing a wall. Two identical Dell laptops running Windows 10 with integrated keyboards and a display diagonal of 14 inches served as workstations. After filling out an informed consent, the students completed the tasks in the following constant order: Simon task, first writing assignment, Color-shape task, second writing assignment, Operation Span task, English placement test, copy task, biographical background questionnaire. The order of the writing tasks (description and argumentation) was balanced between participants. Data collection took approximately 105 minutes. Halfway through the experiment, the participants took a five-minute break. For their participation, they received a monetary compensation.

5. Results

The following section first provides descriptive statistics and correlation matrices of all variables. Subsequently, the results of a principal component analysis are shown, which aimed to summarize and integrate the information of the process-related writing measures. Correlations were calculated to investigate the interrelations between the resulting components, text quality, EF skills, language proficiency, and transcription skills. Lastly, path analyses investigated (1) the interactive effects between EF skills and language proficiency on writing fluency and (2) their indirect effects on text quality. The analyses were performed separately for the argumentative and descriptive writing data since the different requirements of both tasks were expected to produce distinct effects of EF skills on the writing process as a function of language proficiency. All data were pre-processed and analyzed in R (R Core Team, 2021) and are available at: https://www.doi.org/10.6084/ m9.figshare.22730975.

5.1 Descriptive statistics and correlation matrices

Table 2 lists the descriptive statistics of the writing data by task. Table 3 shows their correlations. The descriptive writing data of one participant was missing due to a software failure. Table 4 lists the descriptive statistics of the EF, language proficiency, and transcription measures. The transcription measures of two participants were missing. The data points were replaced using means rather than estimates to avoid boosting statistically significant results. In the subsequent analyses, the response times of the inhibition paradigm were used as an indicator of inhibition skills. The percentage of correct responses could be neglected due to ceiling effects. Response times and the percentage of correct responses of the shifting paradigm were combined via the balanced-integration score (Liesefeld & Janczyk, 2019) to account for accuracy-speed trade-offs (with $r = .28$, $p = .036$), such that higher values relate to better performance, and used as an indicator for shifting skills. All three EF paradigms replicate typical findings, magnitudes, and effect sizes (e.g., Lu & Proctor, 1995; Unsworth et al., 2005). Appendix B provides a detailed description of the analyses of the EF tasks. Table 5 lists the correlations between the EF variables, language proficiency, and transcription skills.

Table 2: Descriptive statistics of all writing measures for the argumentation and description

	Argumentation		Description	
	M	*SD*	*M*	*SD*
Characters per minute (product)	94.95	23.67	117.78	30.91
Characters per minute (process)	114.89	27.21	138.82	34.43
IKI between words (in sec.)	1.09	0.39	0.86	0.32
IKI between sentences (in sec.)	2.93	5.21	2.31	1.68
Revisions per minute	4.95	1.90	5.91	2.38
Characters per burst	26.64	13.45	25.63	8.6

	Argumentation		Description	
	M	*SD*	*M*	*SD*
Product-process ratio	.83	.08	.85	.08

Note. IKI = inter-key interval in seconds

Table 3: Correlations between all writing measures by task

	1	2	3	4	5	6	7
Argumentation							
1 Characters per minute (prod.)							
2 Characters per minute (proc.)	.91***						
3 IKI between words	-.67***	-.77***					
4 IKI between sentences	-.27*	-.32*	.19				
5 Revisions per minute	.22	.50***	-.44***	-.25			
6 Characters per r-burst	.17	-.06	.04	.06	-.73***		
7 Product-process ratio	.27*	-.14	.17	.07	-.66***	.63***	
8 Text quality	.55***	.55***	-.40**	-.24	.18	.12	.08
Description							
1 Characters per minute (prod.)							
2 Characters per minute (proc.)	.92***						
3 IKI between words	-.68***	-.78***					
4 IKI between sentences	-.43***	-.54***	.49***				
5 Revisions per minute	.34*	.59***	-.57***	-.49***			
6 Characters per r-burst	.19	-.04	.21	.20	-.76***		
7 Product-process ratio	.32*	-.06	.20	.23	-.58***	.64***	
8 Text quality	.58***	.52***	-.24	-.18	.10	.29*	.26

Note. with **p* < .05, ***p* < .01, and ****p* < .001; prod. = product; proc = process; IKI = inter-key interval

Table 4: Descriptive statistics of the executive function, language proficiency, and transcription measures

	M	*SD*	*Min.*	*Max.*
Executive functions				
Inhibition (rt)	27.57	15.55	-30.28	64.36
Inhibition (pc)	.96	.03	.87	1
Shifting (rt)	152.76	71.72	15.90	333.68

	M	*SD*	*Min.*	*Max.*
Shifting (pc)	0.92	0.05	0.74	0.99
Updating (pc)	0.52	0.2	0.16	1.00
Language proficiency				
C-test (pc)	0.7	0.17	0.33	0.97
Transcription skills				
Typing speed (IKI in ms)	162.98	27.38	102.00	227.00
Typing accuracy (pc)	.94	.03	.79	.98

Note. rt = reaction time based measure in ms; pc = percentage of correctness; IKI = inter-key interval in ms

Table 5: Correlations between the executive function variables, language proficiency, and transcription skills

Variable	1	2	3	4	5
1 Inhibition					
2 Shifting	-.00				
3 Updating	-.28*	-.02			
4 Language proficiency	-.26*	-.12	.25*		
5 Typing speed	.17	-.20	-.01	-.13	
6 Typing accuracy	-.10	.00	.16	.21	.07

Note. with $^*p < .05$, $^{**}p < .01$, and $^{***}p < .001$

5.2 Preliminary analyses

Principal component analysis

A principal component analysis (PCA; Field, 2012) was conducted to summarize and integrate the information of the set of selected keystroke variables. The correlation matrix indicates high correlations between both production rates (with $r > .91$ for either task condition). Excluding the characters per minute (product) reveals a determinant greater than .00001, a heuristic to avoid extreme multicollinearity (Field et al., 2012, p.770). The remaining data meets the factorability criteria with a Kaiser-Meyer-Olkin measure for sampling adequacy of .65 in the argumentation and .61 in the description, exceeding the recommended value of .61. Bartlett's tests of sphericity are highly significant (argumentation: $\chi^2(15) = 169.01$, $p < .001$; description: $\chi^2(15) = 223.31$, $p < .001$), indicating that the data is sufficiently correlated. For each writing task, an initial PCA results in two components with eigenvalues above Kaiser's criterion of 1 and a clear point of inflection at component two, suggesting a two-component solution. After an oblique rotation (oblimin), production rates and inter-key intervals between words and sentences load on component 1, such that higher values indicate faster writing and shorter pause times. Measures related to revisions load on component 2, such that higher values indicate fewer revisions, longer bursts between revisions, and less revised text. The revision rate also loads slightly but significantly on component 1, reconfirming its close relation to production rates and inter-key intervals. Both components explain 73.1 % of the total variance in the argumentation and 78.8 % in the description. Table 6 shows the summary of the analysis. A loading of ±.50 was used to determine component consistency (Tabachnick & Fidell, 2001). The resulting component scores for each fluency component are labeled productivity and revising, respectively, and used in the subsequent analyses.

Table 6: Summary of the Principal Component Analyses with oblique rotation for the argumentation and description, listing eigenvalues, proportional variance, and cumulative percent of the explained variance for the 2 components in either writing task before rotation and factor loadings after rotation

	Argumentation		Description	
	Component 1	Component 2	Component 1	Component 2
Before rotation				
Eigenvalues	1.59	2.80	3.19	1.54
Proportional variance	.26	.47	.53	.26
Cumulative percent	.26	.73	.53	.79
After rotation				
Characters per minute (process)	**.93**	.00	**.96**	.12
IKI between words	**-.89**	-.01	**-.87**	.05
IKI between sentences	**-.50**	.00	**-.71**	.11
Revisions per minute	.36	**-.79**	.48	**-.70**
Characters per r-burst	.16	**.94**	.08	**.95**
Product-process ratio	.04	**.87**	.08	**.88**
Eigenvalues	2.08	2.30	2.48	2.25
Explained variance	.48	.52	.52	.48

Note. IKI= inter-key interval; factor loadings greater ± .5 are marked in bold.

Table 7. Correlations of the writing components and text quality with executive functions, language proficiency, and transcription skills

	Argumentation			Description		
Variable	Productivity	Revising	Text quality	Productivity	Revising	Text quality
Executive functions						
Inhibition	-.19	.04	-.28*	-.21	.09	-.08
Shifting	.10	-.00	.15	.24*	-.07	.12
Updating	.13	.10	.19	-.11	.15	.05
Language proficiency						
C-test	.22*	.17	.55***	.02	.14	.33**
Transcription skills						
Typing speed	-.52***	.16	-.22*	-.46***	.17	-.18
Typing accuracy	.02	.16	.13	-.01	.25*	.10

Note. with *p < .0 5, **p < .01, and ***p < .001

Correlation analyses

Having aggregated the process-related writing measures into two performance facets for each writing task, the subsequent set of analyses examines the interrelations between the writing performance variables and EF skills, language proficiency, and transcription skills (Table 7). Correlations between most EF skills and either writing process measures or text quality did not reach statistical significance. Enhanced inhibition abilities (as indicated by reduced reaction times) are associated with superior text quality in the argumentation. Additionally, shifting skills are positively correlated with the productivity component in the description. Higher language proficiency is linked to enhanced text quality in both task genres and higher productivity in the argumentative writing data. Furthermore, faster performance in the copy task is associated with higher productivity in either task genre, and higher accuracy relates to more efficient revising in the description.

5.3 Predicting text quality – Does language proficiency moderate the relationship between EF skills and writing fluency?

To examine whether (1) language proficiency moderates the relationship between EF skills and the writing fluency components and (2) the interactive effects indirectly affect text quality, path analyses were conducted using the lavaan-package (Rosseel, 2012) with maximum likelihood estimations, bootstrapping with 5000 resamples, and a 95 % confidence interval (*CI*). The model fit was evaluated with the χ^2-test and global fit indices, i.e., the comparative fit index (CFI), the standardized root mean square residual (SRMR), and a root mean square error approximation (RMSEA). Separate path models were built for each of the EF variables, i.e., inhibition, shifting, and updating. The transcription measures were introduced as co-variates to control for typing skills affecting writing fluency, typing speed for productivity and typing accuracy for revising. All variables were scaled and centered before the analyses to arrive at comparable coefficients.

Shifting

The models have an acceptable fit to the data (argumentation: χ^2 (7) = 8.24, p = .312; CFI = .98, RMSEA = .06, SRMR = .06; description: χ^2 (7) = 5.00, p = .660, CFI = .99, RMSEA = .01, and SRMR = .05). The standardized coefficients of the path models are shown in Figure 1. Overall, the models explain 48 % of the text qualities' variance in the argumentation and 40 % in the description. The productivity component, but not the revising component, explains variances in text quality in the argumentation. In the description, both fluency components contribute to the quality of the final texts. In dependency on language proficiency, shifting has a significant effect on the productivity component in both task genres (argumentation: β = .24, *SE* = .11, *CI* [.06; .47], p = .012; description: β = .25, *SE* = .14, *CI* [.00, .54], p = .046). The interaction significantly affects text quality through productivity in the argumentation (β = .11, *SE* = .06, *CI* [.00, .22], p = .041). In the description, the indirect effect is marginally significant (β = .12, *SE* = .07, *CI* [-.03, .26], *p* = .067). The interaction between shifting and language proficiency significantly affects revising in the description (β = -.36, *SE* = .12, *CI* [-.64, -.17], *p* = .001), which indirectly affects text quality (β = -.12, *SE* = .05, *CI* [-.24, -.04], *p* = .009). In the argumentation, the interaction effect on revising does not reach statistical significance.

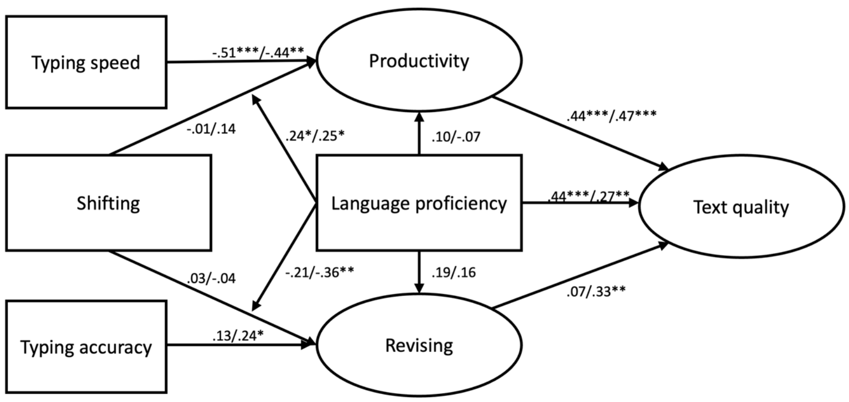

Figure 1: Path model predicting text quality with shifting skills predicting writing fluency (i.e., roductivity and revising) in dependency on language proficiency. Values before the slash represent standardized path coefficients for the argumentation and after the slash for the description with *p < .05, **p < .01, and ***p < .001.

Updating

The models have a good fit to the data (argumentation: χ^2 (7) = 5.22, *p* = .633, CFI = .99, RMSEA = .01, and SRMR = .05; description: χ^2 (7) = 5.07, *p* = .652, CFI = .99, RMSEA = .01, and SRMR = .05). The standardized coefficients for both task genres are shown in Figure 2. Overall, the models explain 48 % of the text qualities' variance in the argumentation and 41 % in the description. The right part of the model including the prediction of text quality with productivity, revising, and language proficiency is largely similar to the shifting model. Updating has a significant effect on the productivity component, but not the revising component, in dependency on language proficiency in the argumentation (β = -.32, *SE* = .10, *CI* [-.54, -.14], *p* = .001). The interaction effect between updating and language proficiency significantly affects text quality through the productivity component (β = -.14, *SE* = .05, *CI* [-.24, -.04], *p* = .004). Updating skills do not contribute to explaining any of the fluency components in the description, and interaction effects do not reach statistical significance.

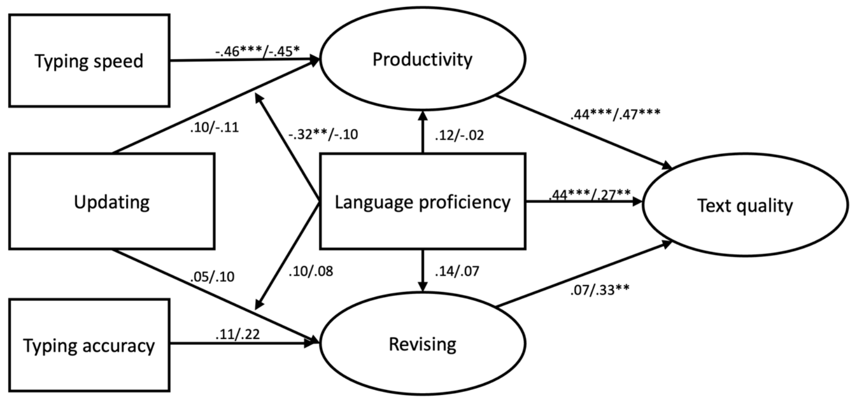

Figure 2: Path model predicting text quality with updating skills predicting writing fluency (i.e., productivity and revising) in dependency on language proficiency. Values before the slash represent standardized path coefficients for the argumentation and after the slash for the description with $^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$.

Inhibition

The model fit the data well (argumentation: χ^2 (7) = 5.90, p = .551, CFI = .98, RMSEA = .01, SRMR = .06; description: χ^2 (7) = 7.78, p = .352, CFI = .98, RMSEA = .04, SRMR = .06). However, inhibition skills do not contribute to explaining either fluency component, and interaction effects do not reach statistical significance in either task genre.

Simple slope analyses

The path analyses revealed four significant interaction effects: in the argumentative writing condition between updating skills and language proficiency as well as shifting skills and language proficiency on the productivity component of writing fluency and in the descriptive writing condition between shifting skills and language proficiency on the productivity and revising component of writing fluency. Simple slope analyses were conducted to determine the direction of the interaction effects (Table 8). In the argumentation, the slopes of updating and shifting with productivity are significant for writers with intermediate and basic as well as proficient language skills (B1-level or below and C1-level or above). In the description, the slopes for shifting with productivity are significant for proficient language skills (C1-level or above). The slopes for shifting and revising are significant for intermediate and basic as well as proficient language users (B1-level or below and C1-level or above). All slopes are non-significant at an upper intermediate language level (B2-level). Figure 3 displays the interaction effects for proficient, upper intermediate, and intermediate language skills (i.e., a C1-, B2-, and B1-level after the CEFR, respectively).

Table 8: Simple slope analyses for the effect of EF skills on writing fluency with language proficiency as a moderator

Task	Interaction	Language proficiency	β	*SE*	*CI*	*P*
Argumentation	Shifting x Productivity	Intermediate	-.19	11	[-.42, .04]	.059
		Upper intermediate	.05	.09	[-.14, .20]	.545
		Proficient	.29	.16	[.03, .60]	.045
	Updating x Productivity	Intermediate	.35	.14	[.07, .63]	.012
		Upper intermediate	.03	.10	[-.18,.22]	.765
		Proficient	-.29	.14	[-.60, -.02]	.040
Description	Shifting x Productivity	Intermediate	-.05	.14	[-.37, .19]	.602
		Upper intermediate	.20	.12	[-.04, .42]	.074
		Proficient	.45	.21	[.01, .85]	.024
	Shifting x Revising	Intermediate	.23	.12	[.07, .54]	.032
		Upper intermediate	-.13	.11	[-.34, .11]	.230
		Proficient	-.48	.20	[-.90, -.13]	.00

Note. Proficient language skills (C1-level with c-test: *Md* + *SD* = .90), upper intermediate language skills (B2-level with c-test: *Md* = .74), and intermediate language skills (B1-level with c-test: *Md* - *SD* = .57) after the Common European Framework of Reference for Languages are taken as a reference.

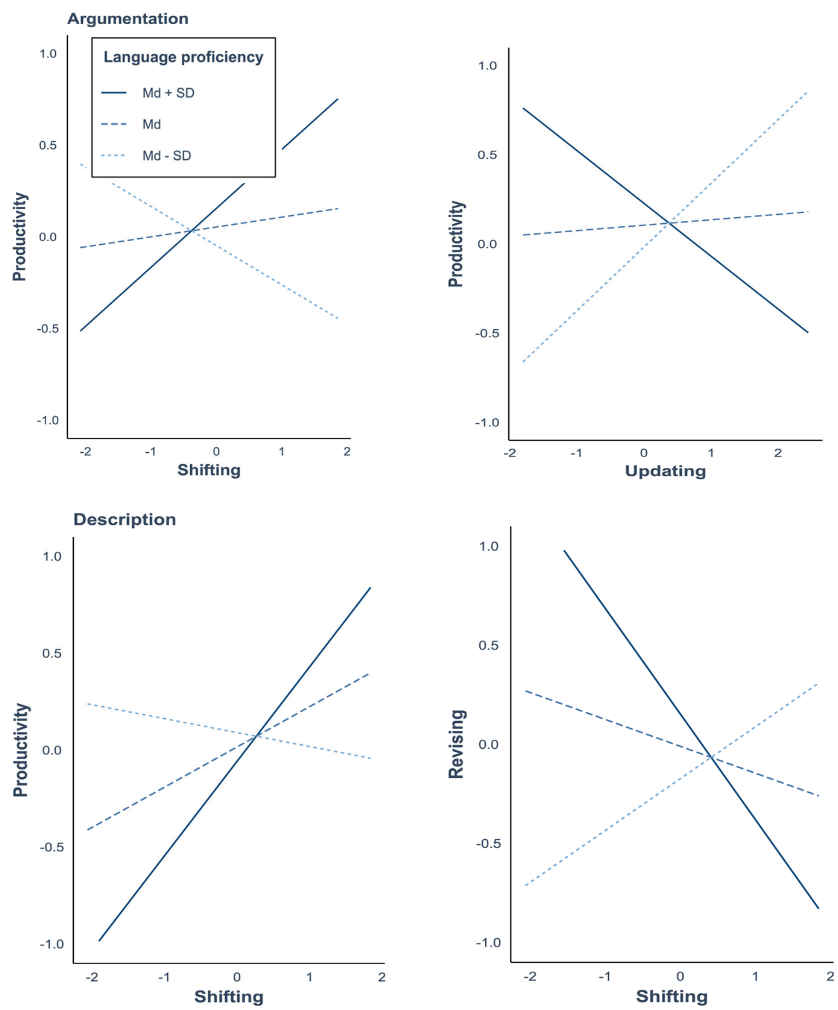

Figure 3: Interaction plots of the relationship between executive function skills and writing fluency with language proficiency as a moderator. Depending on language proficiency, shifting and updating significantly affect productivity in the argumentation (top), and shifting significantly affects productivity and revising in the description (bottom). Proficient language skills (C1-level with c-test: *Md* + *SD* = .90), upper intermediate language skills (B2-level with c-test: *Md* = .74), and intermediate language skills (B1-level with c-test: *Md* - *SD* = .57) after the Common European Framework of Reference for Languages are taken as a reference.

6. Discussion

The present study aimed to investigate the role of language proficiency in the relationship between EF skills and writing performance in L2 text composition. University students composed an argumentative and descriptive text to examine whether interactions between EF skills and language proficiency indirectly affect text quality. Overall, the findings largely confirm the initial hypotheses. Depending on language proficiency, updating and shifting skills, but not inhibition skills, affected the writing process and its fluency, indirectly influencing the quality of the final texts. The results show that the effects of EF skills on writing performance depend on language proficiency and vary with task demands.

6.1 Interactive effects of EF skills and language proficiency on L2 writing processes in argumentative and descriptive writing

Both task genres imposed distinct challenges on writers with different language proficiency levels, reflected in the interactive effects of cognitive and linguistic resources on the writing flow and the indirect effects on the quality of the final product. In the present study, the descriptive writing condition required the participants to recall a familiar and visible object, i.e., their flat. From a linguistic perspective, this is considered a simple task design as it involves knowledge telling and may employ frequent, everyday vocabulary and simple sentence structures. To nevertheless demand EFs, task complexity was increased via the task's procedural demands, i.e., the time to complete the task. The findings indicate that shifting skills played a decisive role in composing a descriptive text under time pressure, reflected in the writers' productivity and revising behavior. Argumentative writing, in contrast, places greater demands on the writer from a cognitive and linguistic perspective as it requires deep processing and deliberation of arguments as well as more sophisticated vocabulary and complex sentence structure. The productivity of writers composing argumentative texts was found to be significantly influenced by their updating and shifting skills, which in turn were dependent on their language proficiency level.

The effects of updating skills on fluent writing processes

Superior updating skills led to significantly higher productivity for writers with basic and intermediate language proficiency in the argumentation. The findings confirm theoretical deliberations that updating skills support manipulating and maintaining information in memory during the formulation process. Further, the results support the idea that non-automated and incomplete linguistic knowledge makes writers more reliant on efficient EFs. Translation processes are typically interrupted - or proceed slower - whenever the information flow at any stage of the writing process is interrupted. This may be caused by non-automated linguistic processes, such as slow lexical retrieval, orthographic encoding, or processing of grammatical structures, or by a non-efficient interaction between processes and resources, such as forgetting the formulated expression before putting it into writing. Superior updating skills seemed to have supported the interactions between cognitive processes and resources, resulting in shorter pauses and faster language production. Writers with inferior updating ability, in contrast, seemed to have needed to stop more often to plan the next item of content or translate ideas into language. The results coincide with findings by Révész et al. (2023), who reported a significant relationship between updating skills and shorter pauses between sentences at the last stages of the composition process.

At advanced language levels, the effects of updating skills on the productivity of writers were reversed, which may be explained by a trade-off between fluent writing processes, on the one hand, and producing complex and accurate language, on the other hand (Skehan, 2009). Fluent text productions are typically associated with high writing performance and text quality. However, disfluency can also improve writing performance, as pauses and revisions may be used for more elaborate writing. A writer might slow down and pause, for instance, to improve the content of their texts, correct the employed language, or generate more diverse and sophisticated language. In the present sample, writers at advanced language levels might have slowed down and paused more often to employ more complex and varied language or elaborate content. The results replicate previous findings by Torres (2023) and Vallejos (2023), who found an association between better updating skills and more frequent and longer pauses. In both studies, the sample comprised highly proficient L2 writers, i.e., emergent and heritage bilinguals, and think-aloud protocols collected by Torres suggest that the writers used the long pauses for linguistic encoding. However, further research is needed to investigate whether longer pauses do indeed relate to higher accuracy and complexity in the final texts for highly proficient L2 writers with good updating capabilities.

In the descriptive writing condition, neither updating skills nor interactions between language proficiency and updating skills significantly affected writing dynamics. Efficient updating processes have been associated with employing complex language in L2 writing (Mavrou, 2020; Vasylets & Marin, 2021). As descriptive writing typically employs simple vocabulary and sentence structures, the findings indicate that efficient updating skills may not be critical for fluent writing.

The effects of shifting skills on fluent writing processes

In both task genres, superior shifting skills were associated with enhanced productivity among writers with advanced language proficiency. The findings are consistent with the theoretical assumption that the ability to efficiently switch between writing sub-processes and tasks allows for more efficient text production with fewer and shorter pauses. At basic and intermediate language proficiency levels, the effects were reversed in the argumentation and non-significant in the description. These results coincide with findings by Révész and colleagues (2023), who reported better shifting skills to relate to longer pausing times in argumentative writing. Depending on skill and knowledge, writers are assumed to use different strategies for text composition (e.g., Torrance & Galbraith, 2005). One could hypothesize that low-proficient writers with better shifting skills might have allocated their resources towards objectives they deemed as pivotal, such as producing more complex language or elaborated content, requiring them to slow down and pause more often in the argumentation. In the description, this may not have been crucial, given the low(er) linguistic demands of the task. Therefore, superior shifting skills did not appear to affect the writers' productivity. Although the simple slope analyses revealed significant effects for writers with high and low language proficiency levels in the argumentation, findings should be regarded with caution. Since only a few participants were observed at this language level, further research, including samples of higher language proficiency, is needed to draw reliable conclusions about the effects of shifting for high- and low-proficient L2 writers.

For intermediate language users, better shifting skills are associated with more efficient revising processes in the description, i.e., they interrupted the writing flow less frequently to perform revisions, produced more characters between revisions, and revised less text overall. The findings are consistent with theoretical assumptions that the evaluative nature of revision processes triggers EFs (Olive, 2021). The effects were reversed for advanced language users, such that writers with superior shifting skills revised more frequently and more texts. One could speculate that writers who are adept at shifting between mental sets will also utilize this skill throughout the writing process, provided that sufficient linguistic resources are available. This may occur even if it is not necessarily conducive to the production of high-quality texts. In accordance with these findings, Révész et al. (2023) observed a greater frequency of within-word pauses among writers with superior shifting skills. The researchers postulated that enhanced shifting abilities may have enabled writers to flexibly transition between lower-level and higher-level writing sub-processes.

In the argumentative writing data, EF skills did not contribute to the explanation of performance variances in the revising behavior. The findings contradict the results by Révész et al. (2017), concatenating superior updating skills to revising behavior in argumentative writing, but are consistent with findings by Révész et al. (2023) and Torres (2023), reporting non-significant effects of EF skills on revising in argumentative writing. The former speculated that revising is an attention-demanding process that is likely more influenced by active strategic behavior than pausing at word boundaries, which is related to linguistic encoding and tends to proceed more automatically with increasing L2 proficiency levels. The different results by Révész and colleagues (2017) might be attributed to methodological differences, as the researchers only analyzed the relationship between two variables at a time. However, frequent revising may result from keyboard unfamiliarity and typing accuracy, which the present study controlled for with the copy task. As there are no comparable studies examining the effects of EFs on revising processes in L2 writing in dependency on language proficiency, more research is needed to determine what accounts for effective revising behavior. The present writing data consisted of two short texts that were composed in a relatively short time window. It would be interesting to investigate, for instance, whether the findings hold in task settings that allow for more extended revision phases.

The (lack of) effects of inhibition skills on fluent writing processes

The effects of inhibition skills on process-related measures did not reach statistical significance in either task genre. The findings contradict previous research, concatenating inhibition processes to pause times (Kim et al., 2021), but are in line with others reporting non-significant effects (Révész et al., 2023). They are surprising at first as, for instance, the ability to efficiently suppress the much more dominant L1 was expected to facilitate lexical recall of complex linguistic structures, which should have resulted in more fluent writing. Since different sub-components of EFs are assumed to interact, resulting in a trade-off between resources devoted to each process, one might infer that updating and shifting are more decisive for the assessed performance facets. Kim and colleagues found better inhibition skills to relate to less and shorter pausing. In contrast to this study, the researchers set the pause threshold to two seconds, which is associated with high-order cognitive writing processes. One could infer that inhibition processes may affect hierarchically high writing processes in particular.

According to Kellogg and colleagues (2013), inhibitory control supports selecting, on the one hand, task-relevant information in the initial planning stages and, on the other hand, appropriate syntax and grammar during translation processes, which may be reflected in the linguistic properties of the resulting text. The correlation analyses revealed a significant relationship between inhibition skills and text quality in the argumentation. The results suggest that good inhibition skills might not affect process- but product-related aspects of writing in the present data. Previous studies, for instance, have linked inhibition skills to the accuracy of texts (Arfé & Danzak, 2020), a variable often used to determine L2 writing performance.

Effects of EF skills, language proficiency, and writing fluency on text quality

The interactions between EFs and language proficiency indirectly affected text quality through process-related writing measures. The findings support the theoretical assumptions that L2 writers with more efficient EF skills and well-developed language skills better cope with the linguistic and structural demands of writing tasks. Language proficiency was the strongest predictor of text quality. The models thus confirm a series of findings in L2 writing studies, demonstrating that well-developed language skills, among others, relate to higher text quality (e.g., Schoonen et al., 2011; Tillema, 2012; Vasylets & Marнn, 2020). Both fluency components contributed to explaining text quality in the description. In the argumentation, however, higher productivity but not efficient revising behavior led to higher quality texts. Text compositions are stop-and-go activities that involve episodes in which language is put into writing (i.e., language bursts) and episodes in which the graphomotor activity is interrupted (i.e., paused). Any writing sub-process may lead to a discontinuity of the writing flow. Planning processes, for instance, may cause pauses when writers must generate new content, or revising processes may lead to interruptions when the writer rereads previous text passages. The findings indicate that while fluent writing processes are associated with high text quality, planning and revision phases may have contributed to more elaborate and sophisticated writing. It is important to note that although the final model explained 45 % of the text qualities' variance, it did not account for more than half of its variance. Other factors, such as content, structure, argumentation, coherence, or linguistic complexity, might have played a decisive role.

6.2 Limitations and outlook

The results indicate a complex relationship between EF skills and writing that depends on language proficiency and varies with task demands. It, hence, seems necessary to factor in writers' language proficiency levels when investigating the influence of EF skills on L2 writing performance. Otherwise, findings will not allow to yield reliable conclusions. The study, however, comprises some limitations that need to be addressed and point to future research. Firstly, the present study used a correlational approach with performance-based cognitive paradigms, which does not allow for causal inferences. However, employing well-established cognitive paradigms designed to engage in one of the sub-domains of EFs provides a highly standardized and reliable way of assessing EFs. The present study used response-time and accuracy analyses that allow for finer-grained distinctions of performance than paradigms comprising more narrow scoring scales. Young adults are at the peak of cognitive development, making it difficult to assess inter-individual differences, particularly in homogeneous groups, such as university students. Using response times and accuracy rates allowed for detecting also small effects of inter-individual differences in EF skills on writing performance. The present sample nevertheless revealed overall relatively efficient executive control processes. It is likely that the effects of inhibition skills on writing fluency, for instance, become detectable in more heterogeneous groups. Secondly, all paradigms were mostly language- and content-free, ensuring a largely unbiased assessment of EFs. However, stimuli must be embedded within a context that allows for deriving reaction times, error rates, or scores to assess the targeted skills. Processing the respective stimuli, hence, necessarily includes systematic variance attributed to non-executive processes, e.g., color, shape, or arithmetical processing (Miyake et al., 2000). Thus, findings must be interpreted in the context of the selected paradigms. Thirdly, the findings are based on one writing sample per genre that provided information about process- and product-related writing characteristics. Since writing performance might vary across assignments, future studies should incorporate multiple data sources to account for the variability between writing tasks. Lastly, the sample comprised students with German as their L1 composing a text in English as an L2. Extending the research to other languages and language backgrounds, educational settings, and socioeconomic backgrounds will lead to a more comprehensive understanding of the role of language proficiency in the relationship between EFs and writing.

Despite its limitations, the study contributes to the limited body of empirical research on the effects of EFs on L2 writing performance, shedding light on the role of language proficiency in the relationship. By extending the analysis to a set of complementary process-related writing measures in two task genres, it could be shown that language proficiency moderates the relationship between EFs and writing on a process level, indirectly affecting writing performance on a product level. Knowing how language proficiency affects the relation between EFs and writing may bring forth useful information for developing interventions that help learners improve marshaling (available) attentional resources and acquire writing strategies tailored to students' language proficiency levels (for a discussion on how to promote EFs during the writing process, see Mason & Brady, 2021). While more research is required, the present results may inform the development of writing interventions that are tailored to language proficiency and cognitive resources, thereby supporting writers in composing texts fluently in an L2.

Notes

[1] Please note that the researchers referred to the construct as attentional capacities. They measured the construct with a Stroop task, which is commonly used to target inhibition skills.

[2] The researcher referred to updating skills as working memory, i.e., the storage and processing functions of working memory, typically measured with complex span tasks, which are also frequently used to assess updating skills (Nyongesa et al., 2019) and have been shown to assess largely similar constructs (e.g., Wilhelm et al., 2013). For consistency, working memory and working memory capacity is referred to as updating, here and in the subsequent paragraphs, if the construct was assessed with complex span tasks, tapping into the updating function of working memory.

References

Abdel Latif, M. M. M. (2013). What do we mean by writing fluency and how can it be validly measured? *Applied Linguistics*, *34*(1), 99–105. https://doi.org/10.1093/applin/ams073

Allport, A. D., & Wylie, G. (2000). Task switching, stimulus-response bindings, and negative priming. *Attention and Performance*, *18*, 33–70.

Arfé, B., & Danzak, R. L. (2020). The influence of first language spelling and response inhibition skills on English-as-an-additional-language spelling. *Cognitive Development*, *56*, 100952. https://doi.org/10.1016/j.cogdev.2020.100952

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: the roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, *100*(1), 320–340. https://doi.org/10.1111/modl.12316

Barkaoui, K. (2019). What can L2 writers' pausing behavior tell us about their L2 writing processes? *Studies in Second Language Acquisition*, *41*(3), 529–554.

https://doi.org/10.1017/S027226311900010X

Becker-Mrotzek, M. (2022). Schreibkompetenz: bberlegungen zu einem didaktischen Konstrukt. In M. Becker-Mrotzek & J. Grabowski (eds.), *Schreibkompetenz in der Sekundarstufe: Theorie, Diagnose und Förderung* (pp. 9–27). Waxmann. https://www.waxmann.com/waxmann-buecher/?tx_p2waxmann_pi2%5bbuchnr%5d=4474&tx_p2waxmann_pi2%5baction%5d=show

Berger, A., & Kiefer, M. (2021). Comparison of different response time outlier exclusion methods: a simulation study. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.675558

Bouwer, R., Lesterhuis, M., De Smedt, F., Van Keer, H., & De Maever, S. (2023). Comparative approaches to the assessment of writing: Reliability and validity of benchmark rating and comparative judgement. *Journal of Writing Research*. Retrieved from https://www.jowr.org/pkp/ojs/index.php/jowr/article/view/867

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: generating text in L1 and L2. *Written Communication*, *18*(1), 80–98. https://doi.org/10.1177/0741088301018001004

Cho, M. (2018). Task complexity, modality, and working memory in L2 task performance. *System*, *72*, 85–98. https://doi.org/10.1016/j.system.2017.10.010

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H.-H. (2019). Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, *41*(3), 583–604. https://doi.org/10.1017/S027226311900007X

Clair-Thompson, H. S., & Wen, Y. (2021). Assessment of executive functions in children. In T. Limpo & T. Olive (Eds.), *Executive Functions and Writing* (pp. 79–102). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0004

Cordeiro, C., Limpo, T., Olive, T., & Castro, S. L. (2020). Do executive functions contribute to writing quality in beginning writers? A longitudinal study with second graders. *Reading and Writing*, *33*(4), 813–833. https://doi.org/10.1007/s11145-019-09963-6

Costa, L.-J. C., Greene, J. A., & Hooper, S. R. (2020). The relations among the development of written language and executive functions for children in elementary school. In R. J. Dippre & T. Phillips (Eds.), *Writing Across the Lifespan* (pp. 159–275). Parlor Press. https://doi.org/10.37514/PER-B.2020.1053.2.10

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press. Retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The contribution of executive functions to narrative writing in fourth grade children. *Reading and Writing*, *28*(7), 989–1011. https://doi.org/10.1007/s11145-015-9558-z

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2017). How executive functions predict development in syntactic complexity of narrative writing in the upper elementary grades. *Reading and Writing*, *30*(1), 209–231. https://doi.org/10.1007/s11145-016-9670-8

Dux Speltz, E., & Chukharev-Hudilainen, E. (2021). The effect of automated fluency-focused feedback on text production. *Journal of Writing Research*, *13*(2), 231–255. https://doi.org/10.17239/jowr-2021.13.02.02

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, *23*(3), 290–325. https://doi.org/10.1191/0265532206lt330oa

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications Ltd.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021). Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – a meta-analysis. *Journal of Cognition*. https://doi.org/10.5334/joc.186

Galbraith, D., & Baaijen, V. M. (2019). Aligning keystrokes with cognitive processes in writing. In E. Lindgren & K. P. H. Sullivan (Eds.), *Observing Writing* (pp. 306–325). Koninklijke Brill NV. https://doi.org/10.1163/9789004392526_015

Grabowski, J. (2008). The internal structure of university student's keyboard skills. *Journal of Writing Research*, *1*(1), 27–52. https://doi.org/10.17239/jowr-2008.01.01.2

Grabowski, J., & Mathiebe, M. (2024). A direct functional measure of text quality: Does the reader understand? *Written Communication*, *41*, 203-229. https://doi.org/10.1177/07410883231222952

Granena, G. (2023). Cognitive individual differences in the process and product of L2 writing. *Studies in Second Language Acquisition*, *45*(3), 765–785. https://doi.org/10.1017/S0272263123000347

Haake, L., Wallot, S., Tschense, M., & Grabowski, J. (2024). Global temporal typing patterns in foreign language writing: exploring language proficiency through recurrence quantification analysis (RQA). *Reading and Writing*, *37*, 385-417. https://doi.org/10.1007/s11145-022-10331-0

Hall, S., Baaijen, V. M., & Galbraith, D. (2022). Constructing theoretically informed measures of pause duration in experimentally manipulated writing. *Reading and Writing*. https://doi.org/10.1007/s11145-022-10284-4

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research*, *55*, 270–279. https://doi.org/10.1007/BF00419687

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

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, Applications* (1st ed., pp. 57–71). Lawrence Erlbaum Associates, Inc.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: a progress report. *Journal of Writing Research*, *5*(2), 159–190. https://doi.org/10.17239/jowr-2013.05.02.1

Kim, M., Tian, Y., & Crossley, S. A. (2021). Exploring the relationships among cognitive and linguistic resources, writing processes, and written products in second language writing. *Journal of Second Language Writing*, *53*, 100824. https://doi.org/10.1016/j.jslw.2021.100824

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*, 390–403. https://doi.org/10.1016/j.jslw.2012.09.003

Kormos, J. (2023). The role of cognitive factors in second language writing and writing to learn a second language. *Studies in Second Language Acquisition*, *45*(3), 622–646. https://doi.org/10.1017/S0272263122000481

Kormos, J., & S6f6r, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism*, *11*(2), 261–271. https://doi.org/10.1017/S1366728908003416

Larigauderie, P., Guignouard, C., & Olive, T. (2020). Proofreading by students: implications of executive and non-executive components of working memory in the detection of phonological, orthographical, and grammatical errors. *Reading and Writing*, *33*(4), 1015–1036. https://doi.org/10.1007/s11145-019-10011-6

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, *30*(3), 358–392. https://doi.org/10.1177/0741088313491692

Leong, C. K., Shum, M. S. K., Tai, C. P., Ki, W. W., & Zhang, D. (2019). Differential contribution of psycholinguistic and cognitive skills to written composition in Chinese as a second language. *Reading and Writing*, *32*(2), 439–466. https://doi.org/10.1007/s11145-018-9873-2

Li, S. (2023). Working memory and second language writing: A systematic review. *Studies in Second Language Acquisition*, 1–33. https://doi.org/10.1017/S0272263123000189

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Limpo, T., & Olive, T. (Eds.). (2021). *Executive Functions and Writing*. Oxford University Press. https://doi.org/10.1093/oso/9780198863564.001.0001

Lindgren, E., Spelman Miller, K., & Sullivan, K. P. H. (2008). Development of fluency and revision in L1 and L2 writing in Swedish high school years eight and nine. *ITL - International Journal of Applied Linguistics*, *156*(1), 133–151. https://doi.org/10.2143/ITL.156.0.2034428

Linnemann, M., Stephany, S., Lemke, V., Bulut, N., Haider, H., Roth, H.-J., & Becker-Mrotzek, M. (2022). The dimensionality of writing and reading fluency and its impact on comprehension and composition. *Journal of Writing Research*, *14*(2), 185–227. https://doi.org/10.17239/jowr-2022.14.02.02

Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2*(2), 174–207. https://doi.org/10.3758/BF03210959

Manchón, R. M., McBride, S., Mellado Martнnez, M. D., & Vasylets, O. (2023). Working memory, L2 proficiency, and task complexity: Independent and interactive effects on L2 written performance. *Studies in Second Language Acquisition*, *45*(3), 737–764. https://doi.org/10.1017/S0272263123000141

Manchón, R. M., Roca De Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing. Implications for theory. In R. Manchón (Ed.), *Writing in Foreign Language Contexts: Learning, Teaching, and Research* (pp. 102–129). Multilingual Matters. https://doi.org/10.21832/9781847691859-008

Mason, L. H., & Brady, S. (2021). Promoting Executive Functions During the Writing Process. In T. Limpo & T. Olive (Eds.), *Executive Functions and Writing* (pp. 181–206). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0008

Mavrou, I. (2020). Working memory, executive functions, and emotional intelligence in second language writing. *Journal of Second Language Writing*, *50*. https://doi.org/10.1016/j.jslw.2020.100758

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, *45*, 31–45. https://doi.org/10.1016/j.jslw.2019.03.002

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Nyongesa, M. K., Ssewanyana, D., Mutua, A. M., Chongwo, E., Scerif, G., Newton, C. R. J. C., & Abubakar, A. (2019). Assessing executive function in adolescence: A scoping review of existing measures and their psychometric robustness. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00311

Olive, T. (2004). Working memory in writing: Empirical evidence from the dual-task technique. *European Psychologist*, *9*(1), 32–42. https://doi.org/10.1027/1016-9040.9.1.32

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, *6*(2), 173–194. https://doi.org/10.17239/jowr-2014.06.02.4

Olive, T. (2021). Executive Functions in Skilled Writers. In T. Limpo & T. Olive (Eds.), *Executive Functions and Writing* (pp. 207–226). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0009

Olive, T. (2022). How do novice and skilled writers engage working memory? In J. W. Schwieter & Z. (Edward)Editors Wen (Eds.), *The Cambridge Handbook of Working Memory and Language* (pp. 504–526). Cambridge University Press. https://doi.org/10.1017/9781108955638.028

Psychology Software Tools, I. (2016). *[E-Prime 3.0]*. https://support.pstnet.com/

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Révész, A., Michel, M., & Lee, M. (2017). *Investigating IELTS Academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory*. The British Council. Retrieved from https://s3.eu-west-2.amazonaws.com/ielts-web-static/production/Research/investigating-ielts-academic-writing-task-2-revesz-et-al-2017.pdf

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behavior: A mixed-methods study. *Studies in Second Language Acquisition*, *41*(3), 605–631. https://doi.org/10.1017/S027226311900024X

Révész, A., Michel, M., & Lee, M. (2023). Exploring the relationship of working memory to the temporal distribution of pausing and revision behaviors during L2 writing. *Studies in Second Language Acquisition*, *45*(3), 680–709. https://doi.org/10.1017/S0272263123000074

Révész, A., Michel, M., Lu, X., Kourtali, N., Lee, M., & Borges, L. (2022). The relationship of proficiency to speed fluency, pausing, and eye-gaze behaviours in L2 writing. *Journal of Second Language Writing*, *58*, 100927. https://doi.org/10.1016/j.jslw.2022.100927

Roca de Larios, J., Manchón, R., Murphy, L., & Marнн, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*(1), 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Rosseel, Y. (2012). {lavaan}: An {R} package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477–1491. https://doi.org/10.1037/0278-7393.31.6.1477

Sicard, V., Moore, R. D., Simard, A., Lavoie, G., & Ellemberg, D. (2022). Psychometric properties of a color-shape version of the switch task. *Applied Neuropsychology: Adult*, *29*(5), 1020–1029. https://doi.org/10.1080/23279095.2020.1842410

Siekmann, L., Parr, J. M., & Busse, V. (2022). Structure and coherence as challenges in composition: A study of assessing less proficient EFL writers' text quality. *Assessing Writing*, *54*, 100672. https://doi.org/10.1016/j.asw.2022.100672

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics*, *6*(1), 99–105. https://doi.org/10.1080/00140136308930679

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Spelman Miller, K., Lindgren, E., & Sullivan, K. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*(3), 433–454. https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Allyn & Bacon.

Tillema, M. (2012). *Writing in first and second language: Empirical studies on text quality and writing processes*. Universiteit Utrecht. Retrieved from

https://dspace.library.uu.nl/handle/1874/241028
Torrance, M., & Galbraith, D. (2005). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 67–82). The Guilford Press.
Torres, J. (2023). Exploring working memory and language dominance in heritage bilinguals' writing processes. *Studies in Second Language Acquisition*, *45*(3), 710–736. https://doi.org/10.1017/S027226312300013X
Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720
Vallejos, C. A. (2020). *Fluency, working memory and second language proficiency in multicompetent writers*. Georgetown University. Retrieved from https://repository.library.georgetown.edu/bitstream/handle/10822/1059667/Vallejos_georgetown_0076D_14678.pdf?sequence=1&isAllowed=y
Van Waes, L., & Leijten, M. (2015). Fluency in writing: a multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition*, *38*, 79–95. https://doi.org/10.1016/j.compcom.2015.09.012
Van Waes, L., Leijten, M., Pauwaert, T., & Van Horenbeeck, E. (2019). A multilingual copy task: measuring typing and motor skills in writing with Inputlog. *Journal of Open Research Software*, *7*(1), 30. https://doi.org/10.5334/jors.234
Van Waes, L., Leijten, M., Roeser, J., Olive, T., & Grabowski, J. (2021). Measuring and assessing typing skills in writing research. *Journal of Writing Research*, *13*(1), 107–153. https://doi.org/10.17239/JOWR-2021.13.01.04
Van Weijen, D. (2008). *Writing processes, text quality, and task effects: Empirical studies in first and second language writing*. Utrecht University. Retrieved from https://dspace.library.uu.nl/handle/1874/33624
Van Weijen, D., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing*, *18*(4), 235–250. https://doi.org/10.1016/j.jslw.2009.06.003
Vandermeulen, N., Lindgren, E., Waldmann, C., & Levlin, M. (2024). Getting a grip on the writing process: (Effective) approaches to write argumentative and narrative texts in L1 and L2. *Journal of Second Language Writing*, *65*, 101113. https://doi.org/10.1016/j.jslw.2024.101113
Vasylets, O., & Marнn, J. (2021). The effects of working memory and L2 proficiency on L2 writing. *Journal of Second Language Writing*, *52*, 100786. https://doi.org/10.1016/j.jslw.2020.100786
Weigle, S. C. (2002). Designing writing assessment tasks. In *Assessing Writing* (pp. 77–107). Cambridge University Press. https://doi.org/10.1017/CBO9780511732997.006
Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00433

Appendix A: Writing Tasks

Argumentative writing assignment

Some politicians raised the following proposition: All short-haul flights should be banned within the European Union. Write a short text of 150 to 250 words in which you explain your opinion on the topic.

Descriptive writing assignment

Describe the flat you currently live in. Write a short text of 150 to 250 words.

Appendix B: Analyses of the executive function skills

The following section provides a detailed description of the design and stimuli used in the executive function paradigms. Subsequently, analyses and results are reported. All tasks were implemented in E-Prime (Psychology Software Tools, 2016); the data were pre-processed and analyzed in R (R Core Team, 2021).

Design and stimuli

Simon task

The Simon task (Simon & Wolf, 1963) is a reaction-time based paradigm that targets inhibition skills. The stimuli comprised red and green smiley faces, which appeared either below, above, on the left, or on the right side of a central fixation cross on the screen. Participants were asked to respond to the color of the stimulus as quickly and accurately as possible, regardless of its location. Thereby, the presented stimuli and response features either matched (the location of the stimulus and response key coincided; compatible condition), they did not match (i.e., the stimulus appeared on the opposite side of the screen as the response key; incompatible condition), or they did not relate (i.e., the stimulus appeared above or below the center of the screen; neutral condition). The Simon task requires inhibiting irrelevant information about the presented stimuli (i.e., the location) and directing attention to task-relevant information (i.e., the color).

Each trial began with a centered fixation cross (+), displayed for 500 ms. The target stimuli followed and remained on the screen until the participant responded, with a maximum of 5000 ms. The task featured five blocks in total. Block 1 was a practice block, introducing the neutral stimuli within eight trials, two for each color and position. Block 2 was an experimental block with neutral stimuli only. Block 3 was again a practice block, introducing compatible and incompatible stimuli. Blocks 4 and 5 were experimental blocks featuring compatible and incompatible trials. Each block comprised two puffer trials, which were excluded from the analysis, and 38 experimental trials, resulting in 38 neutral, 38 compatible, and 38 incompatible trials. A randomized order was used for each participant. The response keys on the keyboard ('q' for green and 'p' for red) were marked accordingly. The participants were advised to use their index fingers and to keep them on the respective response keys.

Color-Shape task

The Color-Shape-Task was modeled closely after Rubin and Meiran (2005) and targets shifting skills. During the reaction time-based task-switching paradigm, participants were asked to respond to either the shape (square/triangle) or color (blue/red) of a stimulus according to a cue that preceded it (a colorful rainbow indicating to respond to the color and a black geometric form indicating to respond to the shape. In the incompatible condition, the task of the previous trial differed from the current (e.g., the current trial required a response to the color, the previous one to the shape) while the tasks matched in the compatible condition.

Each trial started with a fixation cross in the middle of the screen, displayed for 350 ms and followed by the cue for 250 ms. The target stimuli appeared centrally on the screen until the participant responded with a maximum of 5000 ms. The paradigm started with two single-task blocks to familiarize the participants with the target stimuli and response keys. The blocks required the response to either the color (block 1) or the shape (block 2). Both blocks included 16 trials, eight for each color and shape. The following blocks were mixed-task blocks. Block 3 was a practice block to familiarize the participants with the cue and mixed task design. Blocks 4 to 6 featured the experimental blocks with 50 trials each. The first two trials served as puffers, resulting in 72 incompatible (switch) and 72 compatible (non-switch) trials. A single-random order was used for each participant. The participants were instructed to respond as fast and as accurately as possible to the stimuli. They were advised to use the left middle and index fingers to respond to the color of the stimuli (the 'q' for blue and 'w' for red) and the right middle and index fingers to respond to the shape of the stimuli ('o' for triangle and 'p' for circle) and to keep the fingers on the response keys.

Operation Span task

The Operation Span task was closely modeled after Unsworth et al. (2005) and targets updating skills. During the task, participants must maintain sets of unrelated letters in memory while solving a series of mathematical equations. Each letter follows a mathematical operation that is either correct or incorrect (e.g., (3 x 7) + 5 = 26). The operation and letter combinations vary in length from three to seven. Participants were instructed to solve the math problems as quickly and accurately as possible while remembering the presented letters in the correct order. The number of correctly recalled letters is assumed to reflect the skill to store and update information temporarily.

Analyses and results

For the analyses of the reaction time-based Simon and Color-Shape task, practice trials, the first two trials of each block, and trials with reaction times below 200 and above 3000 ms were disregarded. Subsequently, reaction times below or above the mean ±2 ∗ *SD* were defined as outliers for each condition and removed from the analyses, resulting in an average rejection rate of 5.62 % (*SD* = 1.73) in the Simon task and 4.49 % (*SD* = 1.14) in the Color-Shape task. The procedure has been demonstrated to be a reliable outlier exclusion method that reduces Type-I errors (Berger & Kiefer, 2021). The remaining trials were used to calculate the percentage of correct responses (PC) and mean reaction times (RT) per condition. Both reaction time-based tasks replicate typical findings, magnitudes, and effect sizes (e.g., Lu & Proctor, 1995).

In the Simon task, participants responded on average faster in neutral than compatible trials and faster on compatible than incompatible trials (neutral: *M* = 357.67, *SD* = 41.78; compatible: *M* = 386.34, *SD* = 39.43; incompatible: *M* = 412.81, *SD* = 36.15; $F(2, 114) = 111.77$, $p[GG] < .001$, $\eta_g^2 = .25$) with negligible error rates (neutral: *M* = .97, *SD* = .03; compatible: *M* = .97, *SD* = .03; incompatible: *M* = .93, *SD* = .07; $F(2, 114) = 18.01$, *p*[*GG*] $< .001$, $\eta_g^2 = .14$). Two costs were calculated and averaged: the difference in performance (i.e., RTs) between compatible and incompatible trials as an indicator of conflict resolution abilities, and the difference in performance between compatible and neutral trials as an indicator of the cost for having to monitor the presence of a conflict besides the need to solve it (Hommel, 1993).

In the Color-Shape task, participants responded on average faster in the compatible than incompatible condition (compatible: $M = 656.36$; $SD = 148.39$; incompatible: $M = 809.11$, $SD = 161.27$; $F(1, 57) = 263.11$, $p < .001$, $\eta_g^2 = .2$) and more accurately in the compatible than incompatible condition (compatible: $M = .94$; $SD = .05$; incompatible: $M = .91$, $SD = .06$; $F(1, 57) = 27.01$, $p < .001$, $\eta_g^2 = .1$). RTs and PCs were combined via the balanced integration score to account for accuracy-speed trade-offs (with $r = .28$, $p = .036$) as suggested by Liesefeld and Janczyk (2019). The difference in performance between the compatible and incompatible conditions indicated shifting skills.

In the Operation Span Task, the percentage of correctly recalled letters indicated updating skills. As solving the mathematical operations is merely a secondary task, the percentage of correctly solved math operations was neglected.