An investigation into the impact of Google Docs-supported collaborative writing on secondary four students' writing self-regulation, writing self-efficacy and writing performance in a Band 1 school in Hong Kong: a mixed methods study.

Yim Dexter Gareth

MSc in Applied Linguistics for Language Teaching, 2024

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Dexter YIM
Date:	15th September, 2024

Acknowledgement

This dissertation marks a milestone in my life and stands as proof of my perseverance and hard work. I am deeply honored to study at Oxford and would like to thank Dr. Heath Rose for giving me the opportunity to flourish at this top-notch university, meet various experts in Education and Applied Linguistics, and bond with my wonderful classmates.

I would like to express my deepest gratitude to my supervisor, Dr. Elizabeth Wonnacott, who has consistently provided constructive feedback and taught me how to handle statistics – a subject in which I initially felt like an idiot! Her unflagging support and professionalism have been instrumental in my completion of this dissertation.

I would like to thank my good friends in this program – Winny, Benen, Lisa, Filip, Matt, and other classmates who are always good listeners. You are always willing to share your knowledge and lives with me, and I am more than happy to be your listener in the future. I would also like to thank my good friends in Hong Kong (Secondary school: Mr. Tom, Kathy Yip, Wilson Man; University: Kenny Li, Sunny Chan, Tommy Liu, Kenny Yu, Michelle Hau, Beryl Chow, Rebecca Chan, Samson Leung; LKKC: Alex Wong, Victor Wong, Leo Lee, Dian Tai; Teachers and Professors: Melodie Law, Catherine Wong, Jane McKie) who have given me unwavering support and love that still motivate me.

Lastly, and most importantly, to my mum, dad, and brother: thank you for your steadfast belief in my abilities and the sacrifices you have made to see me succeed. Your love and support have always been important to me, especially when I face challenges. Your love has made all the difference in my life. My dad passed away on October 1, 2024, and he could not attend my graduation ceremony. I want to express my deepest respect and gratitude towards him; as his son, I feel immensely proud of him for he is the best dad in the world. Thank you for bringing me into this beautiful world. This dissertation serves as a memory for me and a tribute to my dad, who has been there to witness my growth.

Abstract

Technology-mediated collaborative writing (CW) has been advocated to improve students' second language (L2) proficiency (Li & Chu, 2018; Ma, 2020). However, its impacts on students' writing proficiency, L2 self-regulation (SR), and self-efficacy (SE) are still under-researched in the context of Hong Kong. This study employed both qualitative and quantitative research methods to investigate how Google Docs-mediated CW affects the writing performance, self-regulation, and self-efficacy of Band 1 Secondary Four students in Hong Kong. To this aim, two intact classes with an intermediate English level were selected and assigned to experimental and control groups, each consisting of 15 students. The experimental group received Google Docs-mediated CW instruction and wrote collaboratively, while the control group experienced traditional instruction and wrote individually. Four timed argumentative writing tasks were designed. Both groups were required to write individually in the first and the final writing (the pre-test and post-test); the experimental group worked collaboratively in the second and third writing while the control group did these tasks individually. All essays were scored for complexity, fluency and accuracy. Statistical analyses were conducted on the pre- and post-test essays to explore whether gains were made and whether these differed between the groups. Across the two groups, some areas were found to improve (ratio of clauses per in the ratio of clauses per T-units, morphological complexity (nouns), lexical accuracy, L2 SR, and SE in the WSR test). However there was no evidence that this differed across the experimental and control groups, and thus no evidence of a specific benefit of collaborative writing. Other measures either showed overall no improvement across the groups, or in the case of one measure (morphology errors) a decrease in performance. Moreover, two questionnaires on L2 Writing Self-Regulation Scale (SLWS) and L2 Writing Self-Efficacy Scale (SLWSS) were given to both groups before and after the intervention to measure the changes in their L2 SR and SE. Qualitatively, the collaboratively written assignments of the experimental group, in terms of content, language, and organization, were analyzed using microgenetic methods. Also, semi-structured interviews were also conducted after the intervention and analyzed using thematic analysis. The qualitative data analysis indicated that the students initiated more mediations and showed positive attitudes towards the intervention. Pedagogical suggestions concerning patterns of interaction, the use of the first language (L1), and grouping strategies were offered and discussed.

Keywords: Google Docs-mediated, writing self-efficacy, writing regulation, writing performance, collaborative writing

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE
INTRODUCTION

1. Introduction

The current trend in the field of second language (L2) education has centered around collaborative language learning activities aimed at enhancing the linguistic competencies of L2 students (Oxford, 1997; Swain, 1995). Collaborative writing (CW), which follows Vygotsky's (1978) social-constructivist theory of learning, involves pair-based and group-based tasks that facilitate active participation of students with proficient individuals such as instructors or advanced peers. This helps students to identify and address gaps in their writing performance and co-construct their writing drafts (Rahimi & Fathi, 2022). The emergence of modern educational technologies has reshaped the forms and purposes of writing, which provides new ways to teach and learn second language writing. CW is expected to gain traction due to the proliferation of Web 2.0 technologies, particularly wikis and Google Docs, which are new collaborative writing platforms (Zhang & Chen, 2022). Ortega (2009) contends that, in the current technology-driven world, the integration of computer-mediated activities in language classes is no longer a matter of choice but a necessity. Extensive research has been conducted to explore the potential of computer-mediated communication in language learning, including its capacity to facilitate more negotiations (Blake, 2000; Lee, 2002; Smith, 2005), boost participation (Roed, 2003), influence L2 written outputs (i.e., complexity, accuracy, and fluency (CAF)) (Chen, 2019; Li, 2018; Shehadeh, 2011; Storch, 2005), and promote self-regulation (SR) and self-efficacy (SE) (Rahimi & Fathi, 2022; Su & Zou, 2022; Qiu and Lee, 2020). However, the findings of these studies have been inconsistent, leading to different views on the effectiveness of computer-mediated communication in language learning.

The most recent curriculum guidelines devised by the Education Bureau of Hong Kong stipulate the incorporation of technology-mediated CW as a means of diversifying the learning experiences of students and fostering a cross-pollination of ideas (CDC, 2017). Furthermore, technology-enhanced CW is aligned with one of the learning objectives for senior forms students to cultivate the ability to refine, reorganize, and develop ideas and thereby enhance their expression through independent and collaborative revisions of their written work (CDC, 2017). However, the existing literature on the impact of computer-mediated CW on students' writing performance is limited, as writing is generally viewed as an individual and solitary activity (Storch, 2013) and has been widely accepted among English teachers (Ede & Lunsford, 1990).

When engaging in CW, students' writing performance is influenced by their self-regulation and self-efficacy (Bruning et al., 2013; Pajares & Valiante, 2001; Rahimi & Fathi, 2022; Su & Zou, 2022). However, despite much discussion of SR and SE in L1 writing research, the exploration of these factors in L2 writing has been relatively underexplored (Kormos, 2012; Teng & Zhang, 2018). Furthermore, much of the empirical research on CW has been conducted in university contexts (Yang, 2014; Storch, 2005), with only a small amount of research in L2 secondary school writing classrooms (Villarreal & Gil-Sarratea, 2020; Storch, 2018). There is a dearth of research on the impact of computer-mediated CW on writing performance, L2 SR and SE in the context of English writing classrooms in Hong Kong. Therefore, this study aims to investigate how Google Docs-mediated CW affects the L2 writing self-regulation, writing self-efficacy, and writing performance of Band 1 secondary four students in Hong Kong, thus proffering pedagogical recommendations for teachers, schools, and the Education Bureau.

1.1 Research Questions

This study aims to investigate the impact of computer-mediated CW on the development of writing self-regulation and self-efficacy, as well as writing performance among Hong Kong students, using both qualitative and quantitative research methods. Specifically, this research seeks to answer the following questions:

1. How does a Google Docs-supported collaborative writing approach, with a specific focus on peer feedback and revision, impact the subsequent individual writing performance of Band 1 secondary four students in Hong Kong?

2. How does Google Docs-mediated collaborative writing impact the development of L2 writing self-regulation among Band 1 secondary four students in Hong Kong?

3. To what extent does collaborative writing facilitated by Google Docs enhance L2 writing self-efficacy among Band 1 secondary four students in Hong Kong?

CHAPTER TWO
LITERATURE REVIEW

2. Theoretical framework of Collaborative Writing

2.1 Definition of Collaborative Writing

Collaboration entails shared effort and therefore CW is regarded as a learning activity that involves multiple writers engaging in discussions, negotiating meaning, and making collective decisions at various stages of the writing process to generate a shared text with co-ownership (Li, 2018; Storch, 2013). Specifically, CW not only emphasizes intra-group collaboration, which encompasses brainstorming, outlining, task allocation, joint scaffolding, drafting, revising, and co-editing (Kessler & Bikowski, 2010; Hsu, 2019), but also underscores inter-group interaction, which involves receiving peer feedback from individuals outside the writing group (Elabdali, 2021; Ma, 2020). Moreover, CW is characterized by three principles: (1) frequent communication among group members throughout the writing process, (2) shared responsibility and decision-making, and (3) the creation of a shared document (Ede & Lunsford, 1990; Li, 2018).

Donato (2004) and Kim (2008) posit that collaborative language learning activities involve students with varying abilities who share their linguistic knowledge with other group members to complete the requisite tasks of drafting, discussing ideas, and revising their writing. When they discuss, negotiations often arise during social interactions because of the struggle to attain a shared understanding and expression (Schrage, 1994). Through such interactions, students co-construct their L2 skills, and attain their potential level of functioning (i.e., the Zone of Proximal Development (ZPD)) to effectively self-regulate their cognitive functions (Neumann & McDonough, 2015; Vygotsky, 1986).

In an L2 classroom, the text produced may take the form of a composition or a report, but it also comprises language-focused tasks such as dictogloss, which involves small groups or pairs of students reconstructing a text based on notes taken from a dictated text (Wajnryb, 1990). However, it excludes grammar exercises such as joint editing, cloze, and text reconstruction. In these language-focused tasks, students are not required to generate new text, but instead, they reprocess language that has been previously produced by others (Manchón, 2011). The outcome of CW processes is a text that is co-produced and shared, which cannot be easily attributed to the individual contributions of each writer (Stahl, 2006). Therefore, the co-

written text is deemed as jointly owned, with all writers possessing equal ownership of the created text.

2.2 The occurrence of computer-mediated collaborative writing

The advent of educational technology has led to the evolution of collaborative learning into technology-enhanced collaborative learning (TECL) (Wang, 2015). TECL enhances language input resources, diversifies learning tasks, reduces teachers' lecturing burden, and increases learners' interest (Arnold & Ducate, 2019; Dooly & Sadler, 2016; Tsai, 2019). Among the pedagogies in TECL, computer-mediated collaborative writing refers to the use of technology tools to facilitate synchronous and asynchronous discussions and negotiations during collaborative writing activities in an online environment (Zhang et al., 2021).

Second-generation web applications (Web 2.0) applications encompass podcasts, blogs, wikis, Google Docs, and a variety of social networking sites, such as Facebook and YouTube (Storch, 2013). Among these, wikis and Google Docs are the most widely used online platforms for implementing CW as they facilitate interactive participation in information creation and sharing (Warschauer & Grimes, 2007) and online co-construction of a text in a collaborative and asynchronous manner (Cho & Lim, 2017; Reinhardt, 2019). After defining computer-mediated CW, it is crucial to explore the theoretical underpinnings of CW from various perspectives.

2.3 Socio-constructivist theory (Vygotsky, 1978)

The socio-constructivist theory, as proposed by Vygotsky (1980), emphasizes the active involvement of learners in social interaction as a means of developing knowledge. This theory posits that learners are social beings and that all cognitive development, including language learning, is intricately linked to social interaction. This theory argues that knowledge is co-constructed and negotiated for meaning through constant interactions and assistance between interlocutors or collaborators situated in social activities (Avci & Adiguzel, 2017). To achieve convergence and mutual understanding in cases of communication breakdown, the theory suggests repeating utterances, correcting mistakes, or asking clarification questions (Guo & M€ollering, 2016; Yeh & Chen, 2019). From a sociocultural perspective, CW serves a mutual mediation function (Storch, 2013; Yang, 2014) by allowing learners to pool their linguistic resources, scaffold each other's task performance, and co-construct new knowledge that can

be internalized as language resources for future writing applications (Edstorm, 2015; Vygotsky, 1978).

Collaborative learning writing is informed by various constructs within the sociocultural perspective, such as scaffolding, ZPD, and languaging (collaborative dialoguing and private speech), which will be elaborated on in Sections 2.4-2.6.

2.4 Socio-cognitive theories: interaction hypothesis (Long, 1996) and output hypothesis (Swain, 1995)

As previously stated, CW is a pedagogical method that involves learners' joint participation in writing tasks within pairs or small groups. The two fundamental components of collaborative writing are verbal communication and writing, which are considered socio-cognitive in this context. The significance of these two elements, both from social and cognitive viewpoints, has been explicated by Long's (1983, 1985, 1996) interaction hypothesis and Swain's (1985, 1993, 1995) output hypothesis.

Krashen's comprehensible input hypothesis (1981, 1982, 1985) asserts that for SLA to occur, learners must be exposed to language that they can understand (comprehensible input). The input contains linguistic structures just beyond their current L2 knowledge (i + 1). Long (1983, 1985) supports the significance of comprehensible input for L2 learning and claims that interactional modifications during conversations are a more consistent and prevalent means of achieving comprehensibility. Negotiations for meaning during communication breakdowns often result in modifications of input (e.g., rephrasing), rendering it more comprehensible, and promoting acquisition through the provision of necessary input.

According to Long's interaction hypothesis (1983, 1985, 1996), second language acquisition is facilitated through social interaction, instead of individual learning or instruction. The interactionist perspective on L2 acquisition emphasizes interactionally modified input and output through negotiation of meaning and form, and advocates the use of CW in L2 classrooms in that CW involves learners in discussions that entail negotiation of meaning and form, thereby promoting the acquisition of comprehensible input and production of modified output (Fernandez Dobao, 2012; Storch, 2005).

Swain's (1993, 1995) advocacy for pushed output is rooted in the cognitive processes that are activated during language production. When learners attempt to produce language in

their L2, they may become aware of gaps in their knowledge. This noticing arises when they encounter difficulties in conveying their intended meaning. The identification of gaps in their linguistic knowledge can potentially motivate learners to undertake measures to rectify these deficiencies. These measures may include delving into their internal linguistic resources, testing different hypotheses, analyzing language input more meticulously from a syntactical standpoint rather than a semantic one, and reflecting on their language use (Swain, 1993, 1995). Both theories consider language acquisition as a primarily cognitive process, and thus focus on the factors that trigger learners' internal cognitive processes.

2.5 Swain's notion of languaging and collaborative dialogue

Subsequent to Long's interaction hypothesis, Swain (2000, 2006, 2010) presents a distinct conceptualization and analysis of interaction. Swain's ideas of collaborative dialogue and languaging, which are informed by sociocultural theory, proffer a compelling rationale for tasks that integrate speaking and writing (Storch, 2013). They elucidate how language usage in social interaction serves as a mediator of language learning.

Languaging refers to the use of language as a means to comprehend complex information. When learners tackle challenging tasks, languaging represents the verbal expression of thinking (Swain, 2000, 2006, 2010). This practice is evident across all learning domains and is particularly relevant in the L2 learning domain when learners are faced with linguistic difficulties, such as mastering complex grammatical structures or selecting the most effective way to convey an idea. The significance of languaging lies not only in its problem-solving function but also in its capacity to facilitate the construction of new knowledge and understanding.

To be more specific, languaging comprises two distinct forms, namely private speech and collaborative dialogue (Swain, 2006, 2010), with the former referring to self-directed speech and the latter to speech directed towards others. Collaborative dialogue occurs when learners engage in group or pair work, attempting to resolve a problem through mutual discussion and interaction (Swain & Lapkin, 2002).

Collaborative dialoguing was not always a viable option, particularly when low achievers were engaged in tasks that far exceeded their capabilities. To address this predicament, low achievers resorted to self-regulation via private speech to gain better

command over language structures (Nishioka, 2016). In the context of situated collaborative learning, knowledge co-construction by learners takes place within specific social and physical environments, with learning occurring through communication and the exchange of information (Iba~nez et al., 2011).

When learners collaborate with peers, speech can serve both private and social functions simultaneously (Wells, 1998), offering several additional advantages for language learning. Firstly, overt self-directed verbalization (private speech) in a social context becomes public, enabling others to access and scrutinize the speaker's thought process. Secondly, private speech, such as vocalized hesitations, can denote requests for assistance, prompting listeners to offer timely aid in the form of suggestions, explanations, or confirmations.

CW facilitates engagement in collaborative dialogue (Camps et al., 2000), enabling learners to draw on the linguistic resources of their peers when encountering problems. Collective scaffolding and co-construction of knowledge facilitate language learning (Chang & Wells, 1988; van Lier, 1996) and promote a clearer understanding of learners' own knowledge. Furthermore, interaction with peers exposes learners to diverse viewpoints and can prompt constructive disagreements that require learners to justify their suggestions, enhancing their understanding (Tocalli-Beller, 2003). The sociocultural perspective posits that interaction between learners provides opportunities for both private and collaborative dialogue, which differ from Long's negotiations for meaning.

However, collaborative dialogue does not necessarily have the primary goal of making the conversation or message more comprehensible, nor does it arise purely due to non-comprehension, as suggested by interactionist theorists. Instead, it occurs when learners engage in deliberation, such as determining the optimal means of conveying an idea. This process may involve pooling knowledge sources and co-constructing new language or consolidating existing knowledge, which learners can subsequently appropriate and internalize for their individual use.

2.6 Scaffolding / Collective scaffolding and the Zone of Proximal Development (ZPD)

As previously noted, Vygotsky's sociocultural theory of mind (SCT) is a socio-cognitive framework that underscores the significance of verbal interaction in learning. In this context, interaction is not merely a tool for enhancing input comprehensibility or an avenue for

providing negative feedback and eliciting modified output. Rather, interaction and language usage in interaction are considered pivotal in all facets of cognitive development, including language acquisition.

Lantolf (2005) asserts that social activity and dialogue between individuals are the genesis of all cognitive functions, occurring on the interpersonal or inter-psychological plane, before being internalized as individual resources on the intra-psychological plane. This internalization represents the transition from social knowledge to increasingly individualized knowledge and constitutes cognitive development. In inter-psychological interaction, the form of assistance offered by experts is a critical element, with not all forms leading to effective development. Effective expert assistance must consider the novice's current knowledge state and potential for improvement through assistance (Vygotsky, 1978).

According to Vygotsky (1978), the Zone of Proximal Development (ZPD) denotes the distance between a student's "actual developmental level as determined by independent problem-solving" and their "potential development as determined through problem-solving under adult guidance or collaboration with more capable peers". The ZPD is a collaborative activity, involving both the novice (lower ability students) and the expert (teachers or higher ability students), with the novice's actions or utterances signaling their needs to the expert. Scaffolding, as a form of finely-tuned assistance provided by the expert to the novice, involves encouraging the novice to take greater responsibility for the activity and is mediated by language, which facilitates communication and coordination between the novice and expert (Wells, 1999; Wood et al., 1976).

Collective scaffolding pertains to instances where learners combine their partial knowledge of the L2 to make decisions about vocabulary or linguistic structures. Donato (2004) observed that collective scaffolding allowed his small group learners to perform beyond their current linguistic proficiency levels. However, such scaffolding was primarily observed in groups that functioned as a collective. Donato also noted that the knowledge co-constructed during these interactions was internalized and applied in subsequent independent activities. These findings have prompted Lantolf (2000) to expand the definition of the ZPD to encompass the "collaborative construction of opportunities for learners to develop their mental abilities."

Vygotsky's ZPD and scaffolding can be implemented in a classroom context due to the interactive learning environment, facilitating learning from more capable peers. Lower-achieving students can acquire new knowledge from their higher-achieving counterparts, while the latter can consolidate their understanding by providing elaborated explanations through scaffolding to their lower-achieving peers (Boticki et al., 2011). This approach promotes progress through the ZPD, where learning and development are intertwined (Lee, 2007, 2008), benefiting both high and low achievers.

2.7 Self-regulation and Self-efficacy in the L2 writing context

2.7.1 Self-regulation (SR)

Incorporating collaborative writing in the L2 writing classroom necessitates considering students' L2 writing SR and SE, as previous studies have indicated their positive impacts on L2 writing performance (Boykin, Evmenova, Regan, & Mastropieri, 2019; Su, Li, Liang, & Tsai, 2019; Zimmerman, Bandura, & Martinez-Pons, 1992). Writing tasks typically require "self-planned, self-initiated, and self-sustained" efforts, and the employment of effective self-regulatory strategies can enhance students' writing performance (Zimmerman & Risemberg, 1997).

SR pertains to self-managed thoughts, emotions, and actions that are strategically planned and cyclically adapted to achieve personal goals (Zimmerman, 2000). In collaborative L2 writing, self-regulated writers proactively manage their writing tasks by establishing objectives, addressing writing challenges, and employing and overseeing requisite strategies to enhance their text quality (Csiz.r & Tank., 2017; Fathi, Mohebiniya, & Nourzadeh, 2019; Zimmerman, 2001).

Zimmerman & Risemberg's (1997) social cognitive model of writing highlights environmental, behavioural, and personal writing self-regulation categories. Environmental writing self-regulation pertains to contextual modifications that promote writing task accomplishment, while behavioural writing self-regulation involves self-monitoring, self-evaluation, and self-verbalization of writing processes. Personal writing self-regulation refers to the utilization of cognitive or affective strategies to manage writing tasks. These categories collectively aid writers in achieving desired writing outcomes.

2.7.2 Self-efficacy (SE)

As for L2 writing SE, it is conceptualized as students' confidence and beliefs in their abilities to successfully perform various writing tasks in different domains, including grammar, composition, usage, and mechanical writing skills (Pajares, 2007). In accordance with Bruning et al.'s (2013) framework, L2 writing SE can be subdivided into three elements: ideation, convention, and self-regulation. Ideation denotes a writer's ability to conceptualize and elaborate upon ideas, and skillfully organize them in a coherent and purposeful manner. Convention refers to a writer's proficiency in employing and implementing grammatical guidelines, punctuation usage, and other linguistic conventions. Self-regulation pertains to a writer's ability to set objectives, monitor their progress, and adapt their writing strategies as necessary.

The literature suggests that students' writing SE can positively influence their writing performance (Pajares, 2003; Schunk & Zimmerman, 2007) and writing self-regulation (Bruning et al., 2013; Pajares, 2003). The reason for this is that heightened writing self-efficacy can enhance students' confidence in using self-regulatory writing strategies, resulting in improved writing performance (Su, Li, Liang, & Tsai, 2019; Su et al., 2018; Zimmerman et al., 1992; Zimmerman & Bandura, 1994). There is extensive empirical research indicating a favorable correlation between SE and academic achievement across multiple domains (Honicke & Broadbent, 2016), including the focus of this research – L2 writing.

Having established the relevance of self-regulation and self-efficacy in collaborative writing and other key concepts such as languaging, collaborative dialogue, the ZPD and scaffolding, it is now pertinent to delve into some prominent studies that examine these constructs in greater detail. Through a comprehensive analysis of these studies, we can gain a deeper understanding of the impact of CW on SR and SE and identify effective strategies to enhance these constructs in L2 writing classrooms. Therefore, in the following sections, this paper will review some recent studies that shed light on the impacts of collaborative writing on SR, SE and the writing performance.

2.8 Factors affecting languaging in Collaborative Writing

CW provides learners with numerous opportunities to practice using the target language and enhance their language skills. However, a multitude of factors can influence the learning outcomes of CW. These factors include the type of writing task, the L2 proficiency and

composition of learners, and the relationships that learners form when writing collaboratively (Storch, 2013).

2.8.1 Task type

A variety of writing tasks have been employed in previous studies to explore the effectiveness of CW. These tasks can be broadly categorized into meaning-focused and language-focused tasks (Storch, 2013), as shown in Table 1 below.

To identify the most effective type of task for drawing learners' attention to language, previous research has compared the number and type of Language Related Episodes (LREs), which refer to instances of collaborative dialogue in which learners discuss the language they are producing, question their language use, or self-correct their language production (Swain & Lapkin, 1998), generated by both meaning-focused and language-focused tasks. Researchers have sought to identify patterns in the types of errors or difficulties that learners encounter and the strategies they use to overcome such difficulties. Storch (2013) conducted a review of previous studies (Colina & Mayo, 2007; García Mayo, 2002; Storch, 1998, 2001) and found that controlled grammar tasks such as text reconstruction and editing tend to generate more LREs than meaning-based tasks such as composition and jigsaw due to the nature of the tasks. Meaning-based tasks tend to produce more lexical LREs, which pertain to word meanings and

word choice, while grammar tasks tend to generate more form-based LREs, which relate to different morphosyntactic forms.

2.8.2 The L2 proficiency and composition of learners

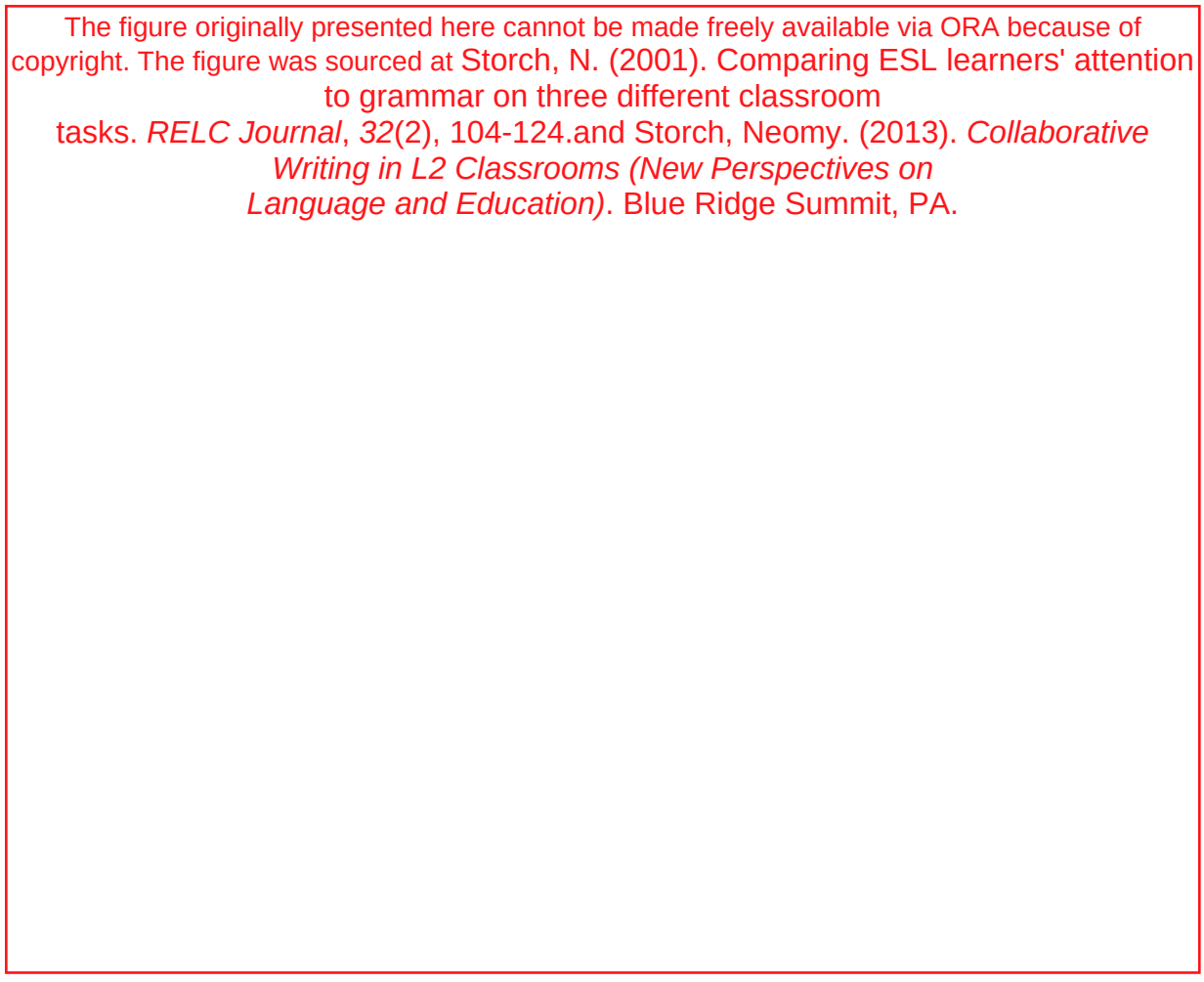
It is inevitable that learners exhibit heterogeneous proficiency levels in L2 classrooms, prompting researchers to investigate the most effective grouping or pairing patterns for language learning in CW. CW involves pairing learners based on their language proficiency levels, creating dyads of either similar proficiency (i.e., high-high pairs and low-low pairs) or different proficiency levels (i.e., high-low pairs) (Niu, Jiang & Deng, 2018). Researchers frequently employ LREs to measure language learning. Typically, researchers quantify lexical and grammatical LREs to assess the effect of proficiency pairing or grouping on language learning in CW, although mixed findings have been reported.

Three studies conducted by Leeser (2004), Kim & McDonough (2008), and Storch & Aldosari (2013) examined the relationship between proficiency pairing, language learning, and LREs in CW among L2 learners. For example, Leeser (2004) investigated 21 dyads and found that higher-proficiency dyads, consisting of higher-proficiency learners only, tended to focus more on grammatical LREs, whereas lower-proficiency learners within mixed and lower-proficiency dyads focused mainly on lexical LREs. However, the study conducted by Kim & McDonough (2008) showed that both higher- and lower-proficiency pairs, regardless of the pairing, mainly concentrated on grammatical LREs. Later, Storch & Aldosari (2013) analyzed 30 learners divided into similar and mixed-L2 proficiency pairs and found that high-high pairs produced the most LREs, followed by high-low and low-low pairs. The study (Storch & Aldosari, 2013) found that all groups, irrespective of their proficiency pairing, mainly dealt with lexical LREs. Notably, the different LREs foci can be attributed to task differences, as Leeser (2004) and Kim & McDonough (2008) used language-focused tasks while Storch & Aldosari (2013) employed meaning-focused tasks. As such, further research is necessary to understand how task types and L2 proficiency affect the number and type of LREs generated, and to identify suitable tasks for different L2 learners with various levels.

2.8.3 The relationships among learners

When considering pairing or grouping students, it is imperative to consider not only their proficiency differences, but also the type of relationship that learners form when working together. It appears that learners with similar levels of proficiency are more inclined to establish

collaborative relationships (except the low-low pairs), compared to pairs with a large proficiency gap (Storch & Aldosari, 2013) because intermediate and advanced learners can initiate more resolutions of LREs and have collective scaffolding, which can affect their groupmates' motivation and engagement (Storch & Aldosari, 2013). Considering this, Storch (2013) devised and illustrated a model of dyadic interaction that encompasses two continua: equality, which is a horizontal continuum that reflects the learners' contribution and control over the task, with high equality indicating not only equal contribution, but also equal decision-making or perceived authority over the task. On the other hand, mutuality is a vertical continuum that reflects the degree of engagement with each other's contribution, with high mutuality indicating interactions that involve co-construction and reciprocal feedback.




The figure originally presented here cannot be made freely available via ORA because of copyright.
The figure was sourced at Storch, N. (2001). Comparing ESL learners' attention to grammar on three different classroom
tasks. *RELC Journal*, *32*(2), 104-124.and Storch, Neomy. (2013). *Collaborative Writing in L2 Classrooms (New Perspectives on
Language and Education)*. Blue Ridge Summit, PA.

Apart from the model of dyadic interaction, there are variables that influence the nature of relationships that learners establish, such as learners' goals (Storch, 2004), their groupmates' praise and motivational phrases (serving as supporters in the affective domain) (Selcuk, Jones & Hana Vonkova, 2021) and their perceptions of their group members' L2 proficiency (Watanabe, 2008). For example, the study (Selcuk, Jones & Vonkova, 2021) recruited 6 EFL learners from a Turkish public high school to participate in CW conducted with their group members in a Facebook group. The six participants with intermediate, pre-intermediate and elementary English levels were divided into two groups consisting of 3 participants each. They were asked to complete a short story writing online outside of school hours in 10 sessions over seven weeks. It employed three data collection methods: group interviews, online written Facebook chats, and Facebook discussion boards. The key findings revealed that the groups elected the group leaders with intermediate English proficiency to facilitate their writing process, and the feedback provided by the leader was found to be useful for the group members' learning. Additionally, based on the findings in the interview, group leaders provided affective support by giving praise and motivational phrases to members, which increased their writing SE and SR in English. However, the interpretation of the findings was limited by the sample size and researcher bias because the researcher played a dual role as an interviewer and facilitator in the intervention.

3. Previous research on collaborative writing in relation to L2 learners' writing performance, self-regulation and self-efficacy

In recent years, a growing body of research has focused on the effects of collaborative writing on L2 learning. After defining the key concepts and theories related to CW, this section aims to synthesize the previous studies that have explored the impacts of collaborative writing on the development of L2 learners' writing skills, self-regulation, and self-efficacy. By examining the findings of existing studies, this review will provide insights into the potential

benefits and challenges of using collaborative writing as a pedagogical tool for L2 writing instruction. Appendix A provides a summary of key features of the studies, including their contexts, tasks, implementation, outcomes, implications, and limitations. A more detailed analysis of the selected studies is provided in Sections 3.1 and 3.2.

3.1 Previous studies presenting CW in relation to L2 learners' writing performance

Dobao's (2012) investigation aimed to examine the impact of CW on the linguistic accuracy of written texts produced by intermediate level university students. The study compared the performance of groups of four learners (n= 15), pairs (n= 15), and individual learners (n= 21) in completing the same jigsaw task. This approach is similar to previous research on jigsaw tasks (e.g., Colina & Mayo, 2007; Storch & Aldosari; Swain & Lapkin, 2002). The study also examined how the number of participants affected CAF of the written texts produced, and the nature of oral interaction between the pairs and groups during the writing process. To analyze the texts produced, the study employed quantitative and qualitative methods. Analysis of the interaction during CW revealed that both groups and pairs devoted a considerable amount of attention to language. Nevertheless, groups produced more LREs and a higher percentage of correctly resolved LREs than pairs and individuals. The study revealed that texts produced by the small groups were the most accurate while those produced by individual writers were the least accurate. However, the difference was not statistically significant. The research concluded that CW could enhance the linguistic accuracy of written texts produced by university students. The results also suggest that groups are more effective than pairs and individuals in producing accurate texts.

Similarly, the study conducted by Wang (2015) investigated the effectiveness of using a collaborative writing platform, wiki, on students' writing performance in an elective course entitled "Business English Writing" at a mid-sized university in Southern Taiwan. The study involved 48 students, divided into two groups: an experimental group (the wiki group) and a control group (the non-wiki group), all taught by the same instructor. The study employed pre-test and post-test writing tasks and a survey questionnaire as research instruments. The results of the pre-test and post-test writings (writing business letters) showed that the wiki group demonstrated greater improvement in their business writing performance compared to the non-wiki group, with a statistically significant difference attributed to the use of wikis in collaborative group writing assignments. An analytical scoring rubric for marking business letters was employed to assess the students' writing. Inter-rater scoring was also used; both the

instructor and the researcher independently rated the writing tests, and a very high consensus and consistency between the two scores was achieved, with an agreement level of 97%. The survey results indicated that students had positive attitudes and perceptions towards wiki-mediated writing and suggested that CW or wiki-mediated CW can enhance students' interest in language learning and improve their writing competencies, which substantiates the findings of Dobao's study (2012) and other previous studies (Malmqvist, 2005; Storch, 2005; Wigglesworth & Storch, 2009), but contradicted some studies that reported no language learning advantage for group over individual task completion when comparing the pre-test and (delayed) post-test performance (Kuiken & Vedder, 2002; Nassaji & Tian, 2010; Reinders, 2009).

3.2 Previous studies presenting CW in relation to L2 learners' self-regulation and self-efficacy

Previous studies focused on how CW impacted learners' writing performance but only a few studies have explored L2 learners' SR and SE in the CW context. To address this gap, the study (Qiu & Lee, 2020) aimed to investigate the effects of collaborative writing on the development of writing skills, regulated learning and self-efficacy beliefs of 24 EFL university students in China in the context of a semester-long English course. The participants were intermediate-level English learners. The study collected and analyzed the learners' individual writing before and after CW, task discussions, reflective notes, and follow-up interviews. The results of the study demonstrated that learners' writing skills, in terms of CFA, improved after the CW intervention. Furthermore, the total length of all 18 audio recordings (each including the discussion of a CW task among the participants was analyzed and the study found that the learners engaged in higher frequencies of co-regulation, self-regulation, and socially shared regulation monitoring episodes during the second and third CW tasks compared to the initial task discussion. Analysis of the learners' self-reports also indicated their heightened awareness of self-efficacy and self-regulation for L2 writing. These findings provide insights into the teaching of L2 writing, suggesting that teachers may benefit from implementing collaborative writing activities to enhance learners' writing skills and encourage them to regulate their own writing process.

The study (Rahimi & Fathi, 2021) utilized a sequential explanatory mixed-methods to investigate the effect of wiki-mediated collaborative writing on the writing performance, writing self-regulation, and writing self-efficacy of English as foreign language (EFL) students

in Iran. Oxford Placement Test (OPT) (Allan, 2004) was employed to ensure that students' English proficiency was at the same level. Two intact classes comprising 67 EFL students were randomly assigned to an experimental group (n= 35) and a control group (n= 32) for 12 ninety-minute sessions. The experimental group received wiki-mediated collaborative writing instruction, while the control group experienced face-to-face (non-wiki) collaborative writing instruction. The collaborative setting in this study was different from most of the CW studies (Dobao 2012; Qiu & Lee, 2020). It aimed to explore the impact of implementing CW on a Wiki-mediated platform on participants' writing proficiency, L2 SR and SE. In contrast, other studies typically focused on an experimental group engaging in CW while the control group wrote individually. This distinction was crucial as it highlighted the variable of technological mediation and its potential influence on the learning outcomes of CW. The study collected quantitative and qualitative data using two timed writing tasks, SLWS and SLWSS (Han & Hiver, 2018), and an individual semi-structured interview. The types and frequencies of the students' writing mediations in the wiki-mediated collaborative writing group were also identified. Paired samples t-tests and one-way ANCOVA were used to analyze the quantitative data, while language-related episodes and thematic analyses were applied to analyze the qualitative data. Their collaboratively written samples in the pre- and post-tests were marked based on the rubric developed by Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981). The rubric includes a 100-point scheme: content, 25 points; organisation, 25 points; grammar, 25 points; vocabulary, 15 points; and writing mechanics, 10 points. The results indicated that there were significant differences between the wiki-mediated and non-wiki collaborative writing groups in improving the students' writing performance, L2 writing SR and SE, suggesting that the wiki-mediated CW group was more effective than the non-wiki CW group in developing students' writing performance, self-regulation and self-efficacy. Moreover, the qualitative data analyses (microgenetic method and individual semi-structured interview) also further confirmed that the EFL Iranian students in the wiki-mediated collaborative writing group benefited from various peer writing mediations that contributed to their writing content, organization, and language use, and SRSE in the wiki space.

CHAPTER THREE
METHODOLOGY

4. Research Gap and Rationale for the Current Study

Despite the extensive research on CW in L2 learning, there remains a research gap in the context of Hong Kong. Firstly, the existing literature lacks studies on technology-mediated CW conducted specifically in the Hong Kong context. Secondly, while previous research has delved into various writing tasks, there is a notable absence of studies focusing on meaning-focused genres like argumentative writing, which is a part of the curriculum of the DSE exam (CDC, 2017). Thirdly, most of the participants in the previous studies were university students, and little research has focused on the local Band 1 secondary school students in the Hong Kong context. Fourthly, research exploring the relationship between CW and students' self-efficacy and self-regulation is scarce in the Hong Kong setting. Moreover, the existing research typically isolates specific components of CAF, rather than providing a holistic examination. The omission of any one of these components – complexity, accuracy, or fluency – can result in overgeneralization in research outcomes. For example, prior studies (Storch, 2005; Fernández Dobao, 2012) have contributed insights into syntactic complexity, lexical complexity and fluency but overlook morphological complexity. Additionally, there is an ambiguity in the term "accuracy" with studies not clarifying which linguistic dimensions it encompasses. Therefore, this study seeks to address the above gaps by exploring the impacts of the technology-mediated CW intervention on the students' CAF in the argumentative writing, L2 self-efficacy and self-regulation.

5. Research design and data collection

5.1 Participants

In this pilot study conducted over two months, a cohort of 30 Form 4 students, aged 15 to 16, were selected. These participants are native speakers of Cantonese and have acquired English as their second language. It is common practice for classrooms in Hong Kong to be categorized into subgroups of 15 students each to cater to the diverse learning needs of students. Therefore, the composition of the 30 students, distributed equally between a control group and an experimental group. The experimental group was divided into five smaller groups, with each subgroup comprising three students. Both the control and experimental groups were the two

intact average groups of this school and their English proficiency was similar, based on the school results.

Selecting 30 participants (normally a teacher teaches three groups of students in Hong Kong) could strike a balance between gathering a substantial amount of data and managing the logistical challenges related to data collection and analysis. The students were specifically chosen because they had not previously been exposed to collaborative writing instruction before the intervention.

5.2 Tasks design

Students were asked to complete four written assignments as part of their coursework in the second term. Two writing tasks were taken from the Hong Kong Diploma of Secondary Education (HKDSE) exam and the other two were taken from the school-based writing topics. Students were required to write argumentative essays of 400 words. Among all the topics in the HKDSE exams, only two argumentative writing topics fall under the theme of social media. It is pivotal to choose topics that are closely related to control for the variable of topic familiarity because it can ensure that any variation in the quality of argumentative writing is due to students' L2 writing skills rather than their prior knowledge of the topic. The topics of the four writing tasks were listed in Table 3 below:

Writing	Year	Theme	Topic
1	2024	Social media	There is an ongoing debate about whether social media enhances or undermines social skills. While some argue that social media helps students develop interpersonal skills and maintain relationships, others contend that it replaces face-to-face interaction, potentially leading to poorer social skills and addiction. Write an argumentative writing. Choose one side of the argument and state your position, giving reasons and evidence to support your view.

2	HKDSE 2019	Social media	The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Hong Kong Diploma of Secondary Education (HKDSE) English Language past papers
3	HKDSE 2021	Social media	
4	2024	Social media	Social media companies collect vast amounts of user data, raising concerns about privacy violations and data misuse. This debate centers on the balance between user privacy and the commercial interests of social media companies. Write an argumentative essay. • Discuss whether social media platforms have a responsibility to protect user privacy, or should users be solely responsible for managing their own data. • Give reasons to support your point of view.

Table 3: Writing topics of the intervention

Both groups were required to write individually in writing 1 and 4 as the pre-test and post-test. In writing 2 and 3, the experimental group was asked to write collaboratively and the instructions were provided by the researcher. Instructional processes were shown below in the collaborative writing group:

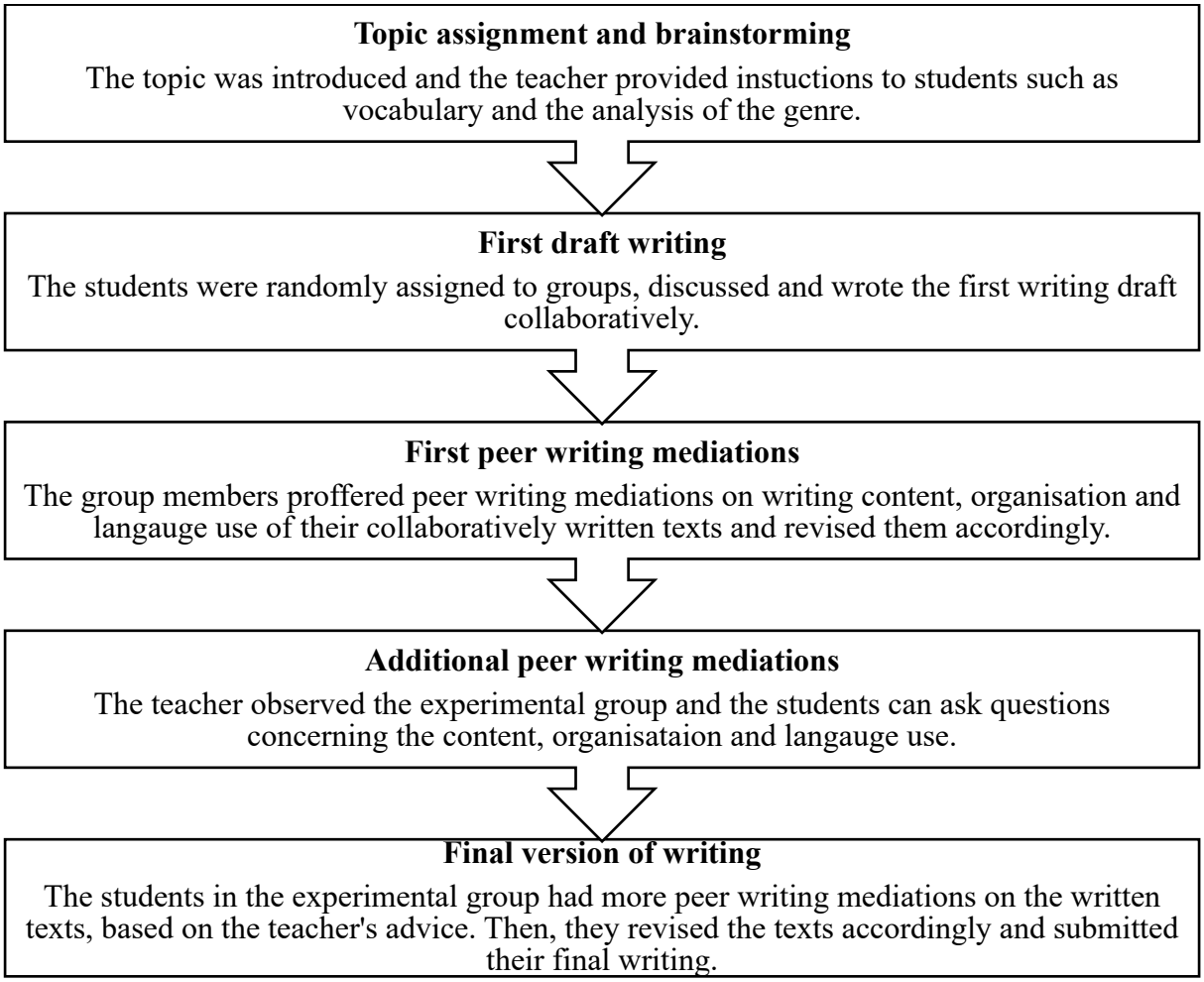


Figure 2: Instructional processes followed in the CW 2 and 3

5.3 Procedures

Based on the school curriculum, three lessons were provided for each writing assignment, with each lesson lasting 40 minutes. Before the students wrote, the teacher spent 50 minutes analyzing the text-type, the topic and formality, teaching thematic vocabulary related to social media, and interacting with the students. Then, they were given 70 minutes to write about the topics. For the pre-test, both groups needed to write individually. For writing topics 2 and 3, in the experimental group, 15 students were divided into five small groups. Each group had their own iPads and worked together on the shared Google Docs. They discussed

the topics with each other and completed the writing task collaboratively. As for the control group, students worked on the same tasks at the same time as the experimental group, but they worked individually. They also finished the task using Google Docs. In order to aid later analysis of the collaborative process, the students were asked to use the tracking and commenting functions while they were revising others' writing. After the intervention, they were required to complete writing 4 as the post-test.

Both the experimental and control groups received a link to a Microsoft From and answer two questionnaires containing items from SLWS and SLWSS devised by Han & Hiver (2018) before they worked on writing 1. They competed this form in the classroom before the pre-test in mid-April to assess their L2 writing SR and SE. After the intervention in mid-June, they completed the same online questionnaire (via Microsoft forms) again and the four students from the experimental group were selected for the semi-structured interview to detect change in self-regulation and self-efficacy. The procedures and the timeline of the intervention were depicted in the flow diagram presented below:

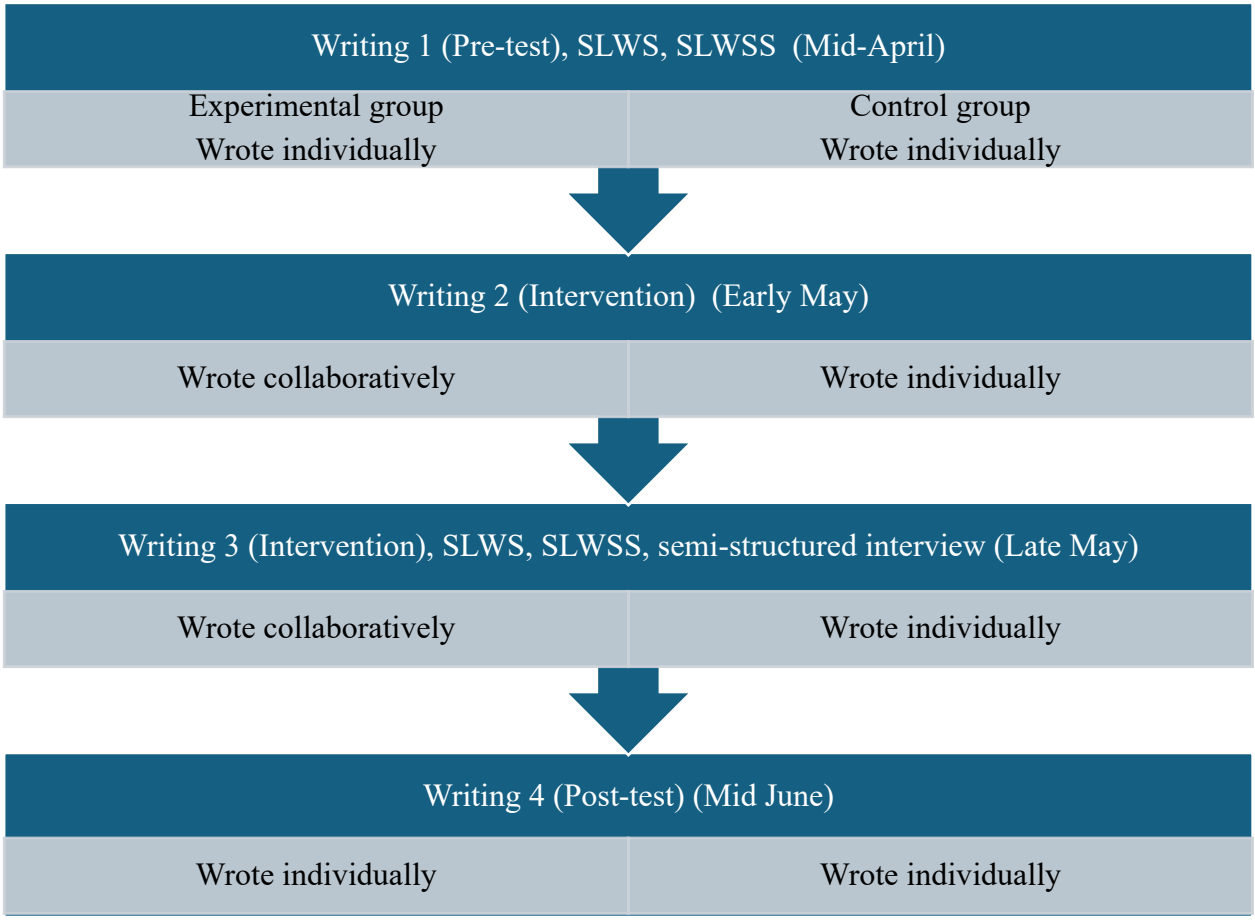


Figure 3: The procedures and the timeline of the intervention

5.4 Measurement of L2 writing

The collaboratively written documents of writing 2 and writing 3 were analysed using the microgenetic method (Lantolf, 2000) and measured in terms of complexity, accuracy, and fluency (CAF) (Skehan 2014; Wigglesworth & Storch, 2009).

5.4.1 Microgenetic methods

The study employed a qualitative microgenetic approach to gain a deeper and more detailed understanding of the findings. To investigate the students' collaborative processes of writing argumentative texts, Language-Related Episodes (LREs) on the shared Google Docs were recorded and analyzed. An analysis of LREs involved an examination of their immediate interactions, including comments, editing, and revisions on the Google Docs (Rahimi & Fathi, 2022). After they finished their drafts, they could revise their collaboratively written works by referring to the comments on the content and organization, and addressing issues related to language use. To more accurately analyze the LREs, the research studied the record of the tracked changes and commented functions used by the students. The researcher closely studied the changes in their writings and evaluated their mediations in relation to three distinct domains: the content of their writing, the organization of their ideas, and their use of language, as delineated in Figure 4. A rubric (Table 4), modified with reference to Figure 4, has been employed to quantitatively measure these domains.

5.4.2 CAF measurement

The measurement of CAF (Complexity Accuracy and Fluency) model facilitates a holistic evaluation by separating linguistic performance into three distinct but interrelated components. In the exploration of CAF, it is important to acknowledge the diversity of measurement approaches adopted by different researchers (Pallotti, 2021). For instance, to measure complexity, Skehan's research (2014) explored syntactic complexity, which has been operationalized using T-units, sentences, and clauses. On the other hand, studies conducted by Avci & Adiguzel (2017), Lee (2004) and Liu et al. (2014) investigated the impacts of CW on the students' lexical complexity. To provide a holistic evaluation, in this study, syntactic, lexical and morphological complexity were measures as follows:

Complexity: In this study, syntactic, lexical and morphological complexity were measured; first, to measure syntactic complexity, the number of T-units, total number of clauses, independent clauses and dependent clauses, the ratio of clauses per T-units and the number of words per clause were calculated (Pallotti, 2021).

Regarding lexical complexity, LancsLex, a lexical analysis tool (Brezina & Gablasova, 2017) evaluates the lexical coverage of texts by comparing them to the New General Service List (NGSL) (Brezina & Gablasova, 2015), which comprises the 2,490 most frequent words in English. The non-GSL is a list of words that tend to be less frequent and more sophisticated. This tool can identify the 2,500 most common English words in the text, distinguishes specialized vocabulary, and filters out proper nouns and numbers. By employing grammatical analysis, LancsLex can differentiate between words with identical forms but belonging to different word classes, such as "walk" as a noun or verb. Moreover, it can also identify lemmas, grouping together all instances of the same word regardless of morphological variations, including singular and plural forms (Brezina & Gablasova, 2015). Specifically, this lexical analysis tool can generate data on the total number of words in the text and the percentage of the text covered by the NGSL and provide a breakdown of the lexical items found in the NGSL into specific frequency-based categories, such as 0-500, 501-1000, 1001-2500.

Lastly, Morpho complexity tool (Brezina & Gablasova, 2015), was also used to calculate Morphological Complexity Index (Pallotti, 2015) of verbs and nouns. The analysis focused on the morphological complexity of verbs and nouns in the four writing tasks. Higher scores indicate a higher level of morphological complexity, while lower scores suggest a lower level of complexity.

Accuracy: It refers to how well linguistic performance aligns with the norms of the target language and has been operationalized through various measures, such as the average number of errors per unit (clause, sentence, T-Unit) or, conversely, by the proportion of error-free units (Pallotti, 2021). In this study, it calculated the ratio of error-free clauses, and counted the number of syntactic, morphological and lexical errors to gauge accuracy.

Fluency: Most of the research on fluency has focused on oral productions. Usually, speed fluency, breakdown fluency and repair fluency are measured in speaking, but in writing, according to Pallotti (2021), assessing L2 written productions is more difficult. A gross indicator may be the count of words or characters per minute, assuming that the duration of writing for each composition is recorded Pallotti (2021). Therefore, fluency is measured by calculating words per minute.

5.5 Measurement of L2 writing self-regulation and self-efficacy

The Second Language Writing Self-Regulation Scale (SLWS) (see Appendix B), developed by Han and Hiver (2018), assesses students' L2 writing self-regulation. Adapted from Tseng, Dörnyei, and Schmitt's work (2006), the 9-item SLWS measured EFL students' strategies for planning, organizing, and managing their L2 writing-specific goals and processes. The 5-point Likert scale questionnaire, with responses ranging from 1 (never) to 5 (always), evaluated the extent of self-regulation. The scores for the five-point Likert-scale items are interpreted as follows: 1 is never, 2 is rare, 3 is sometimes, 4 is often and 5 is always.

Similarly, the Second Language Writing Self-Efficacy Scale (SLWSS) (see Appendix C) devised by Han and Hiver (2018) gauges students' L2 writing self-efficacy. The 7-item SLWSS, adapted from Mills, Pajares, and Herron's instrument (2006), examined students' beliefs and confidence in their L2 writing abilities. Similar to SLWS, the SLWSS utilized a 5-point Likert scale with 1 representing never and 5 representing always to assess students'

writing self-efficacy. The scores for the five-point Likert-scale items are interpreted as follows: 1 is never, 2 is rare, 3 is sometimes, 4 is often and 5 is always.

5.6 Semi-structured interviews

Individual semi-structured interviews were conducted immediately after the CW 3 with four members of the experimental group to gain qualitative insights into their self-regulation strategies and self-efficacy in the collaborative writing lessons. They were selected based on the researcher's observation during the intervention. They were from different groups and were allowed to bring their collaboratively written works to jog their memory about the processes in CW. These interviews were conducted in the participants' L2, English, because the school is an EMI school and English-related activities must be conducted in English. Each interview took approximately 25 to 30 minutes. Subsequent to the interviews, the audio recordings were transcribed to facilitate detailed data analysis. Thematic analysis, based on Rahimi & Fathi (2022), was modified to find out the positive and negative perceptions of CW (Tables 5 and 6) in relation to their self-regulation, self-efficacy and their writing abilities To enhance the reliability of the responses, the interview questions (see appendix D) were expounded upon during each session as necessary, ensuring the participants' understanding and the precision of their answers (Rahimi & Fathi, 2022).

CHAPTER FOUR
FINDINGS

6. Results

The results of this study are presented in three sections, each focusing on a different aspect of the investigation. Sections 6.1 and 6.2 evaluate the writing performance of the participants in four writing tasks. This analysis employs two complementary approaches: the microgenetic method and CAF measurement. Section 6.3 delves into the data gathered from the SLWS, SLWSS and the semi-structured interviews with thematic analysis. This analysis aims to further elucidate the findings from the previous sections to proffer a more holistic understanding of the participants' writing performance, L2 SE and SR.

6.1 Microgenetic method

For the two collaborative writing tasks, due to the limited number of writing samples collected in writing 2 and 3, this study would not conduct a statistical analysis using SPSS; instead, the microgenetic method was adopted and the claims in this section were interpreted based on ten collaboratively written samples. The microgenetic method analyzed the collaborative writing mediations of the experimental group and LREs of their collaborative writings were carried out to explore their moments, editing and revisions (i.e. their peer writing mediations) on collaborative writing tasks regarding writing content, organization and language use (Rahimi & Fathi, 2022). Each category is further divided into specific sub-categories, and the frequency of each sub-category is recorded for both writing tasks. The frequency of the mediations on these three areas were counted and compared. If a sentence contained two language-related mistakes, they were counted twice respectively in each sub-category. Table 4 presents a comparative analysis of peer writing mediations of two writing tasks 2 and 3 for the experimental group.

Experimental Group	**Writing 2**	**Writing 3**
Content		
Clarity of message	15	22
Topic development / elaboration (provide specific details)	2	9

Relevance of message	1	0
Synthesis of information	1	4
The total frequency (Content)	19	35
Language		
Word choice	46	37
Active / Passive	1	2
Adverbs	0	1
Auxiliary verbs	1	1
Singular / plural	1	0
Articles	2	9
Demonstratives	1	0
Prepositions	6	2
Pronouns	2	1
Tenses	5	6
Subject-verb agreement	4	2
Punctuation	4	1
Parts of speech	3	1
Spelling	1	4
Word order	3	2
Relative Clauses	5	1
Capitalization	1	1
Superlatives / Comparatives	0	0
To-infinitives and gerunds	4	1
The total frequency (Language)	90	72
Organization		
Formality (use of diction, abbreviation, contraction)	4	8

The use of connectives and conjunctions	3	4
The total frequency (Organization)	7	12
Total number of mediations	116	119

Table 4: Peer writing mediations of writing 2 and 3 for the experimental group

6.1.1 Content-related peer writing mediations

As evident in Table 4 above, the experimental group had a considerable tendency towards addressing the content-related issues such as clarity of message. Overall, their content-related mediations increased substantially from 19 times in writing 2 to 35 times in writing 3. The most frequent technology-mediated CW mediation in both writing 2 and 3 was clarity of message (15 times and 22 times, respectively). This suggests that the experimental group mostly revised the messages to make them clearer and easier to understand by deleting some unnecessary or confusing information, or further elaborating and explaining the ideas.

In the writing content category, the students also proffered more topic development / elaboration mediations. The number of the topic developments also increased significantly from 2 times to 9 times, which suggests that the experimental group deemed the purpose of the written text (topic development) important and therefore they placed more emphasis on elaboration to enhance the clarity and depth of their ideas. The bolded parts in the examples below were the changes, or the additions made by the students. For example, concerning topic development / elaboration and clarity of message, the student in group B added the university name to make it more specific and convincing. Moreover, the original sentence "more than fifty percent has been scammed" was grammatically incorrect and incomplete, and the peer-edited version clarified who is being scammed and added "manipulate their content for financial purposes" to enrich the content, which accurately describes the deceptive actions.

Example 1 (Group B, Writing 3)	According to a research conducted by ~~**a university**~~ ***(Stanford University)*** in the United States, more than fifty percent ~~**has been scammed by social media advertisement as people may use social media influencers and edit them to earn profit**~~ of respondents reported being scammed by social media advertisements as people may utilize social media influencers to ~~**earn more money**~~ ***manipulate their content for financial purposes***.

In example 2, the students also elaborated on the idea by providing concrete evidence like the organization (the Pew Research Center) and the statistics (41%) to show that the claim was based on real-world data rather than their mere opinion.

Example 2 (Group A, Writing 2)	It often results in more aggressive exchanges and behaviour, rather than respectful and rational conversation. ***For instance, a study by the Pew Research Center found that 41% of Americans have been personally subjected to harassing or abusive behavior online, showing how anonymity can exacerbate negative interactions.***

It is normal to express writing purposes in the opening of argumentative writing. However, the original sentence written by group E in writing 2 sounds more like a conclusion. The addition "Therefore, this essay will argue that social media is stifling rational debate online and explore the implications of this trend" improves topic development and clarity of message by explicitly stating the main argument and the writing purpose.

Example 3 (Group E, Writing 2)	Due to the ~~**anonymity**~~ ***anonymous*** nature of social media, the quality of speech made by internet users ~~**could be**~~ ***cannot be*** guaranteed. Under such circumstances, rational discussions are usually ~~held back~~ ***stifled*** by ***the*** misuse of social media~~***. I would conclude that social media is stifling rational debate online.***~~ ***Therefore, this essay will argue that social media is stifling rational debate online and explore the implications of this trend.***

6.1.2 Language-related peer writing mediations

As for language use, the total number of language-related peer writing mediations was 90 in writing 2, while it decreased to 72 in writing 3. The most frequent language-related mediation in both writing tasks was word choice, with 83 occurrences in writing 2 and 3. Group B paid attention to formality and therefore chose to use a synonym of "buy". The synonym "purchase" was used because "buy" is more casual and widely used in everyday conversation, whereas "purchase" conveys a more formal tone in argumentative writing.

Example 1 (Group B, Writing 3)	Some even hire popular celebrities to encourage people to ~~**buy**~~ ***purchase*** their products to earn money.

Another common language-related mediation was related to articles. There was a notable increase in articles from 2 instances in writing 2 to 9 in writing 3. The definite article "the" was deleted in the original sentences because it was redundant. "Students" and "social media" are general nouns in this topic sentence. Moreover, the group failed to identify another grammatical mistake – a comma splice, which occurs when two independent clauses are joined by a comma.

Example 2 (Group A, Writing 2)	Social media platforms also stifle public debate to some extent, ~~**the**~~ students use their own anonymous accounts to spread some ideas on ~~**the**~~ social media platforms.

The last example of language-related mediation involves the use of modal verbs. Group E changed the gerund "falling" to "fall" because the modal verb "may" should be followed by the base form of the verb. One of the group members used the comment function to provide more examples of modal verbs, such as might, can, will, would, should, must, and shall, to remind the student to use the base form after these modal verbs.

Example 3 (Group E, Writing 3)	Unsuspecting users may ~~**falling**~~ ***fall*** victim to these scams, leading to financial losses and a loss of trust in online advertising as a whole. (Comment: ***You should use the base form of the verb after these modal verbs like might, can, will, would, should, must, and shall.)***

6.1.3 Organization-related peer writing mediations

Comparatively speaking, the total number of the organization-related mediations was significantly lower than that of content- and language-related mediations. However, there was an increase from 7 times in writing 2 to 12 in writing 3. The most prominent sub-category in this area was formality (diction, abbreviation, contraction), with 4 instances in writing 2 and 8 in writing 3. The results suggested that the participants began to pay more attention to the organizational aspects of their writing in writing 3, despite this area still receiving less focus

compared to content and language. In argumentative writing 2 and 3, it is necessary to maintain a formal tone to persuade the audience and sound more credible and convincing.

The original sentence included unnecessary words like “still the same”, which can be considered as redundant. Moreover, the original sentence had an informal tone, and the revised sentence had a more formal tone.

Example 1 (Group B, writing 3)	It is still the same problem. → ***The problem still exists***.

In example 2, the student changed “I’m writing to” to “will” to avoid the use of a contraction because contraction is generally deemed informal and more suitable for casual writing such as letter of advice, diary writing and informal emails.

Example 2 (Group D, Writing 2)	I’~~**m writing to**~~ ***will*** analyze and explore not only the credibility of advertising companies…

6.1.4 Unsuccessful or unnecessary writing mediations

Nonetheless, not all peer writing mediations regarding the three categories resulted in successful revisions. As illustrated in the following examples, there were some unnecessary changes or even incorrect mediations regarding content. In example 1, there were two changes; the student in Group A changed “82%” to “over 80%” and “disclose any history of personal use” to “disclose any experience of using the products.” The first revision was reasonable because rounding numbers up or down is a common practice when we report them. However, the second change was considered unnecessary because it did not enhance clarity or specificity. Additionally, the revised version did not introduce new or additional information, making the change redundant.

Example 1 (Group A, Writing 3)	Moreover, ~~**82%**~~ ***over 80%*** of respondents said social media influencers should absolutely ~~**disclose any history of personal use**~~ ***disclose any experience of using the products*** when promoting a product.

Furthermore, some language-mediations were incorrect, especially grammatical mistakes. Some of the examples were summarized below. In example 1, the change was grammatically incorrect as the sentence lacked a main verb, resulting in a fragment sentence rather than a complete sentence. However, this student was able to replace "mean" with "refer to", which is a more formal word choice.

Example 1 (Group B, writing 3)	Traditional media advertisements typically ~~mean~~ ***referring to*** advertising through televisions, newspapers and presses.

Another unnecessary change was related to parts of speech. In the original sentence, "proved" indicated that social media has demonstrated the ability to facilitate interactions, but the peer changed a verb to a noun, which does not fit the context of the sentence.

Example 2 (Group C, Writing 2)	In conclusion, social media has been ~~**proved**~~ ***proof*** to facilitate interactions with other people.

In example 3, in the original sentence, "as a result" is an idiomatic expression and was used properly, but the student changed "a" to "the", which is not correct as it sounds less natural without a particular context mentioned in that paragraph.

Example 3 (Group D, Writing 3)	As ~~**a**~~ ***the*** result, they have gained trust and confidence from their followers.

Apart from grammatical mistakes, some unnecessary changes were also related to the use of vocabulary such as word choice and synonyms. The changes were deemed unnecessary in example 1. Firstly, the phrase "to conclude" is a formal way to draw the arguments to a close. It is appropriate for summarizing the main arguments. On the other hand, "in summary" is also a phrase used to encapsulate the main arguments, but it lacks the definitive closure that "to conclude" provides. Secondly, the term "significant" in the original sentence conveys that the concerns are important, while "pivotal" means crucial. Since their meanings are similar, the change was unnecessary.

Example 1 (Group E, Writing 3)	~~**To conclude**~~ ***In summary***, although social media influencers have surged in popularity as a marketing tool, ~~**significant**~~ ***pivotal*** concerns remain regarding their trustworthiness when compared to traditional media.

This section has qualitatively examined the CW mediations of the experimental group using the microgenetic method to analyze their peer writing revisions across content, language, and organization. The findings suggest that the mediations, in terms of content, tend to improve clarity and topic development, as seen in the above examples. For language, effective revisions included enhancing formality by choosing precise words, such as changing "buy" to "purchase". As for organization, improvements were made by maintaining a formal tone by avoiding using a contraction. However, the above improvements were based on the limited number of collaboratively written samples, and not all peer mediations were successful or necessary, as some changes created other grammatical errors.

6.2 CAF Measurement

This section outlines the methodology and findings from tests conducted to gauge syntactic complexity, morphological complexity, lexical complexity, accuracy, and fluency across control and experimental groups. Tests for normality indicated that some of the variables were not normally distributed. Therefore, for consistency, non-parametric tests were applied throughout.

6.2.1 Syntactic complexity

To examine the syntactic complexity of the control and experimental groups, four sets of measures on syntactic complexity were examined, including the number of T-units, total number of clauses, the ratio of clauses per T-unit, and the number of words per clause (Lu, 2010).

Concerning the number of T-units, numerically, as seen in Figure 5, the mean number of t-units increased in the experimental group (pre-test mean = 15.66, post-test mean= 18) and decreased in the control group (pre-test mean= 22, post-test mean=19.47). However, an independent-samples Mann-Whitney U test (MWU test) comparing the gains scores from the two conditions found that this difference was not significant (U=149.00, z=1.523, p=.137). A further one sample Wilcoxon signed rank test (WSR test) found that - across the two conditions - the gains scores did not differ from zero (W=129.5, z = -5.87, p=.557), indicating no evidence of an overall difference from pre- to post test.

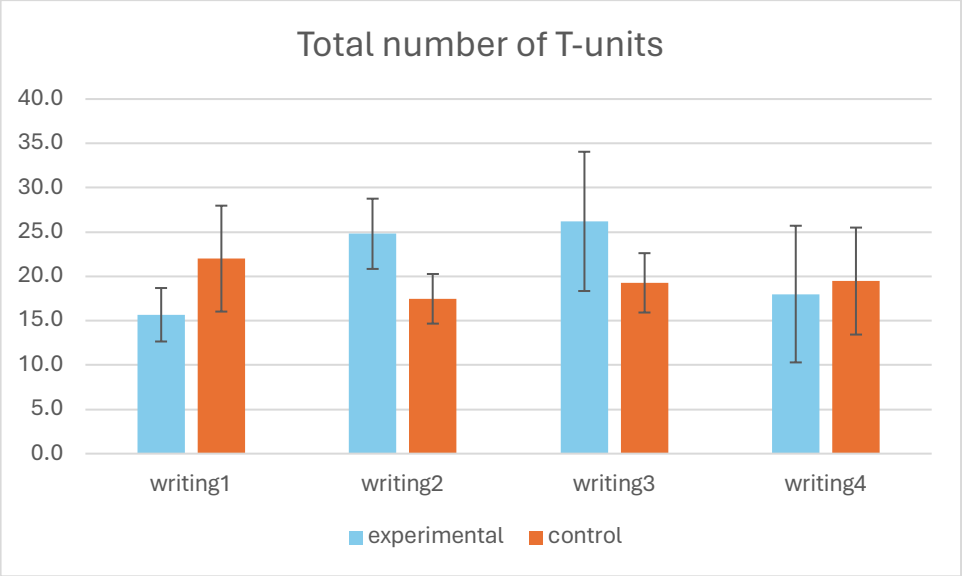


Figure 5: Bar graph showing the mean number of T-units for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

As for the number of clauses, the mean number of clauses increased numerically in both groups, as shown in Figure 6; the experimental group demonstrated an improvement in using more clauses in their writing, with the mean score increasing from 30.2 in the pre-test to 31.4 in the post-test. Comparatively, the control group showed a more substantial improvement, as evidenced by the increased mean score from 28.93 to 33.33. However, an MWU test conducted to compare the gains in the number of clauses used from the pre- to the post-test for the

experimental and control groups revealed no significant difference between the groups (U = 92.5, z = -0.833, p = .412). An WSR test found a non-significant difference in overall increase from pre to post test (W = 321, z = 1.823, p = .068), indicating no evidence of change from the pre-test to the post-test across the conditions.

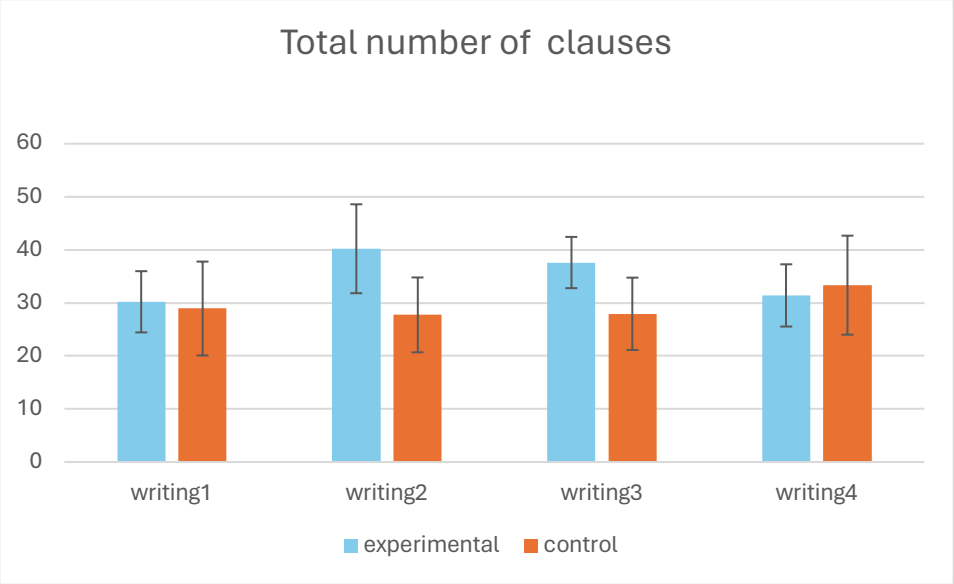


Figure 6: Bar graph showing the mean number of clauses for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

In terms of the ratio of clauses per T-units, Figure 7 showed that the ratio of clauses per T-units differed numerically between the experimental and control groups in the pre-test and post-test. The mean ratio of clauses per T-units in the experimental group slightly decreased from 1.94 in the pre-test to 1.87 in the post-test. Conversely, there was an increase in the mean ratio of clauses per T-units in the control group from 1.35 in the pre-test to 1.75 in the post test. The MWU test comparing the gains in the ratio of clauses per t-units between the two conditions revealed a significant difference between the experimental and control groups (U=38.5, z=-3.071, p=.001). As evidence was found that the groups were different, the WSR test on the gain scores was conducted separately for each of the control and experimental groups. For the experimental group, the results (W= 47.5, z= -0.710, p= 0.478) indicated no evidence of an overall difference between the pre-test and post-test. In contrast, the test results for the control group (W=117.5, z=3.267, p=.001) indicated evidence of an overall difference between the pre-test and post-test.

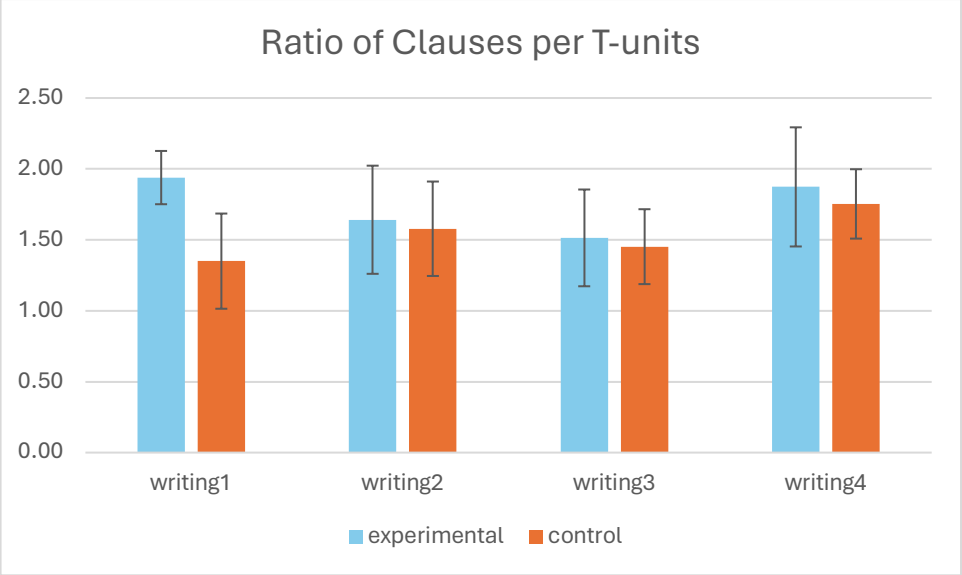


Figure 7: Bar graph showing the mean ratio of clauses per T-units for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

Finally, to measure syntactic complexity, this study also investigated the change in the number of words per clause of both groups. Figure 8 showed that the mean number of words per clause increased numerically from 14.12 in the pre-test to 14.41 in the post-test in the experimental group while the mean number of words per clause dropped significantly from 14.10 in the pre-test to 11.91 in the post-test in the control group. Despite the seemingly positive impact of CW on the experimental group, the MWU test found that the difference between the gains in the two groups was not statistically significant (U=147, z=1.431, p=.161). Moreover, the WSR test also found that the gains scores did not differ significantly from zero (W=203.5, z=-.597, p=.551), indicating no evidence of a significant change from pre- to post-test looking across the groups.

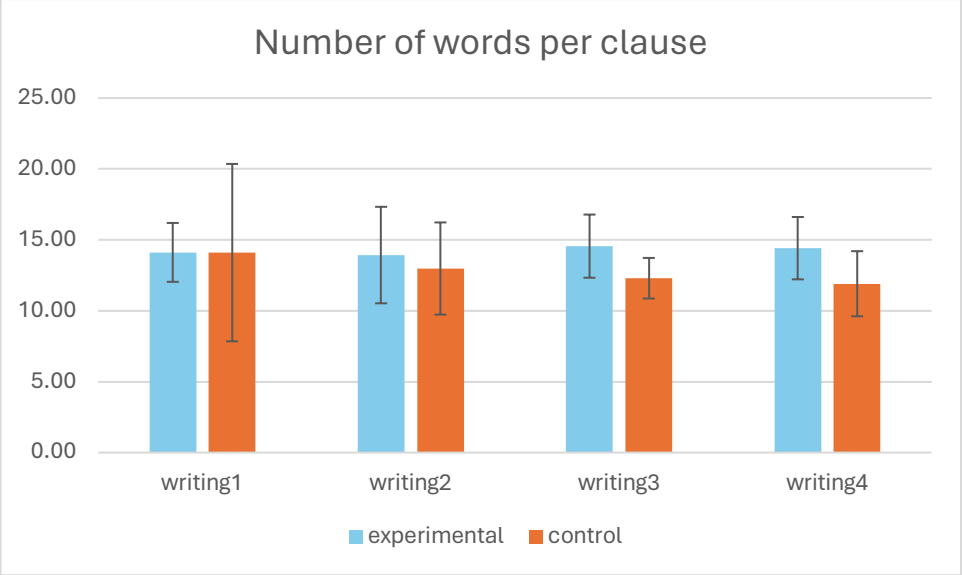


Figure 8: Bar graph showing the mean number of words per clause for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

6.2.2 Lexical complexity

This study investigated the impacts of the CW intervention on lexical complexity in the compositions. It analyzed the use of new GSL 500/1000/2500 words and non-GSL words in the compositions to measure the changes in lexical complexity between the experimental and control groups. Figure 9 showed that the means and standard deviation of NGSL and non-NGSL words. Both the experimental and control groups, in terms of lexical complexity, showed numerically an increase in the use of NGSL from the pre-test to the post-test. The experimental group had the mean NGSL coverage of 0.88 in the pre-test and 0.89 in the post-test. On the other hand, the control group showed numerically the mean NGSL coverage of 0.85 in the pre-test and 0.88 in the post-test. Nonetheless, the MWU test revealed that the difference was not statistically significant (U=80.5, z=-1.332, p=.187). Furthermore, the WSR test also indicated that the gains scores across both conditions did not significantly differ from zero (W=270.5, z=1.540, p=.123), indicating no evidence of an overall difference from pre-test to post-test.

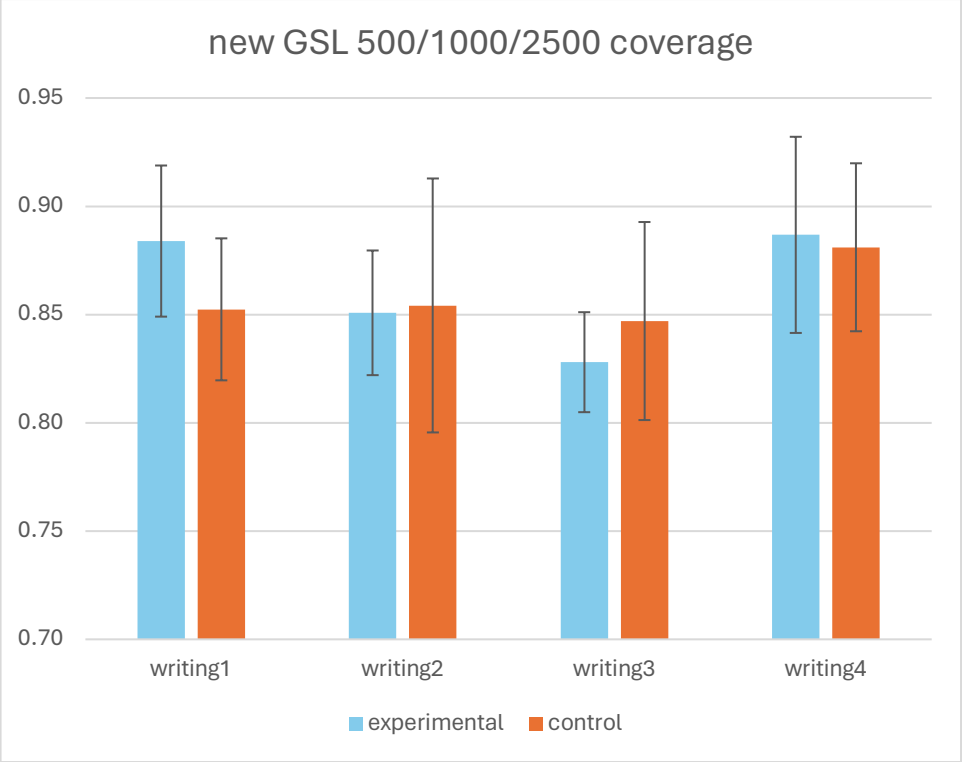


Figure 9: Bar graph showing mean new GSL 500/1000/2500 words for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

As for the non-GSL words, Figure 10 summarized the means and standard deviation of the non-GSL words in the compositions of both groups. Compared to the control group (pre-test mean=0.15, post-test mean=0.12), the mean non-GSL words in the experimental group only decreased slightly from 0.12 to 0.11. To assess the statistical significance of the differences between the experimental and control groups, in terms of gains in the use of non-GSL words, the MWU test was conducted. The results (U=144.5, z=1.332, p=.187) indicated that the differences in gains were not statistically significant across the two conditions. As for the WSR test, the results also showed that the gains scores did not differ significantly from zero (W=135.5, z=-1.54, p=.123), suggesting no significant evidence of an overall difference in using the non-GSL words.

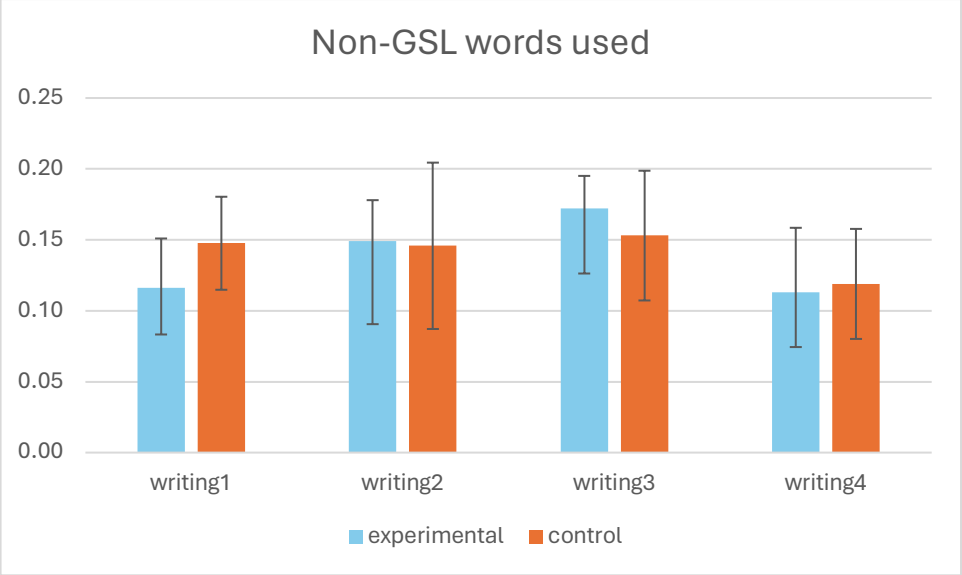


Figure 10: Bar graph showing mean non-GSL words for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

6.2.3 Morphological complexity

Figure 11 below showed that the mean MC (verbs) of the experimental group numerically decreased from 4.52 in the pre-test to 4.44 in the post-test, but that of the control group numerically increased from 4.5 to 4.67. The results of the MWU (U=99, z=-.560, p=.595) suggested that the differences in MC (verbs) between the two groups were not statistically significant. As for the WSR test, the outcomes indicated that the gains scores did not differ significantly from zero (W=246, z=.278, p=.781), suggesting no evidence of significant improvements in MC (verbs) during the intervention.

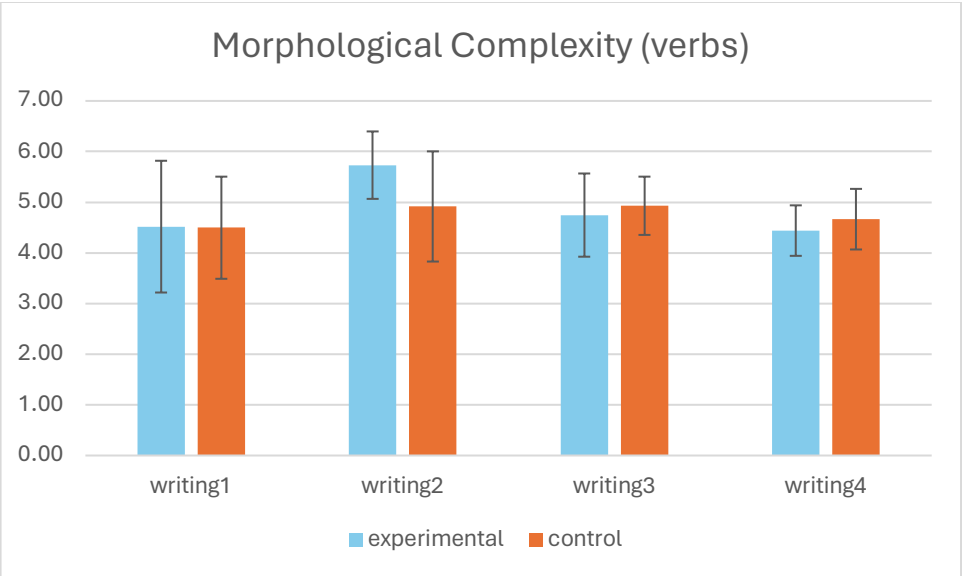


Figure 11: Bar graph showing mean MC (verbs) for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

On the other hand, in terms of MC (nouns), a numerical increase was observed in both groups. Figure 12 showed that the mean MC (nouns) numerically increased from 2.06 in the pre-test and 2.6 in the post-test and from 2.07 and 2.35 in the experimental and control groups, respectively. The results of the MWU test (U=144, z=1.307, p=.202) indicated that the differences in MC between the two groups were not statistically significant. However, the results of WSR test (W=420, z=3.857, $p < .001$) demonstrated that the gains scores differed from zero, implying the significant improvement in MC (nouns) across the two conditions, from the pre-test to the post-test.

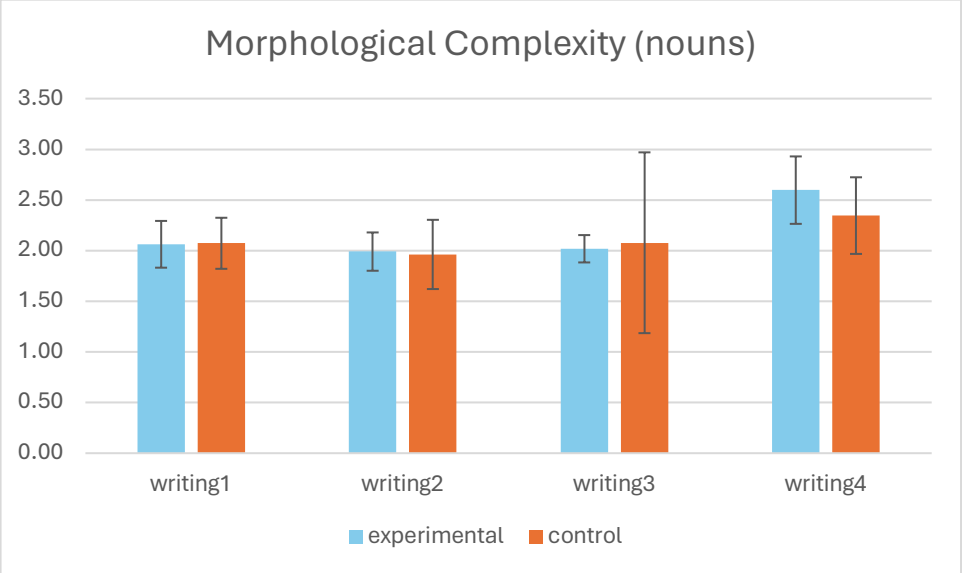


Figure 12: Bar graph showing mean MC (nouns) for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

In summary, this study found that the intervention did not lead to a significant change in MC (verbs), but the MWU test indicated that the differences in MC (nouns) between the two groups were not statistically significant. This suggested that the intervention, in terms of MC (nouns), did not lead to significant differences between the two groups. However, the WSR test found that there were significant improvements in MC (nouns) looking across both groups.

6.3 Accuracy

To evaluate the writing accuracy of the experimental and control groups, the ratio of error-free clauses, and the number of syntactic, morphological, and lexical errors, were analyzed in four writing tasks. Figure 13 presented a summary of the mean and standard deviation of the ratio of error-free clauses. A numerical increase from 0.63 in the pre-test to 0.71 in the post-test in the experimental group and from 0.56 to 0.64 in the control group was observed. However, the results of the MWU test (U=100, z=-.519, p=.624) indicated that the differences in the changes in the ratio of error-free clauses between the two groups were not statistically significant. Similarly, the WSR test showed that the gains scores did not differ significantly from zero (W=307, z=1.533, p=.125), and therefore there was no evidence of an overall difference in the ratio of error-free clauses across the two conditions, from the pre-test to the post-test.

Figure 13: Bar graph showing the mean ratio of error-free clauses for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

As for syntactic errors in the compositions, the means and standard deviations for syntactic errors in the experimental and control groups were listed in Figure 14 The mean syntactic errors made by the students decreased numerically from 5.27 in the pre-test to 4.27 in the post- test in the experimental group and 6.53 to 6.13 in the control group, respectively.

The results of the MWU test (U=109.5, z=-.125, p=.902) indicated that the differences in the changes in syntactic errors between the two groups were not statistically significant, suggesting that the intervention did not significantly impact the syntactic accuracy of the student writings when comparing the two groups. Additionally, the outcomes of the WSR test (W=163.5, z=-.614, p=.539) showed that the changes in syntactic errors across both groups did not differ significantly from zero, indicating no evidence of an overall improvement in syntactic accuracy from the pre- to the post-test in either group.

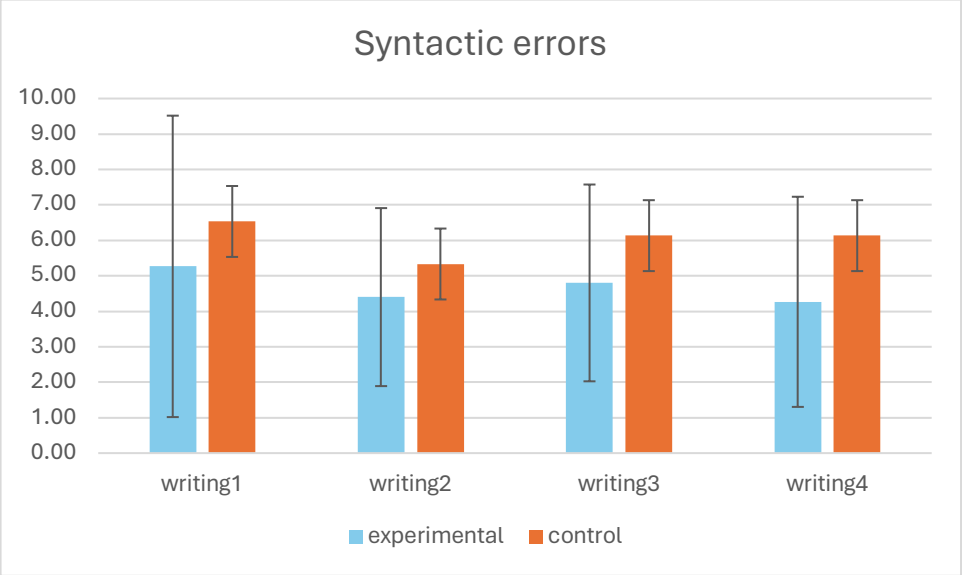


Figure 14: Bar graph showing mean syntactic errors for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

Numerically, Figure 15 showed that both the experimental group and the control displayed the increased mean morphological error from 2.07 to 3.60, and 2.87 to 4.33, respectively. The results of the MWU test (U=121.5, z=.376, p=.713) suggested no significant difference between the two groups, indicating that the intervention did not statistically enhance morphological accuracy in comparison to the control group. On the other hand, the results of the WSR test (W=279, z=2.170, p=.030) showed that the gains scores differed from zero, indicating evidence of a statistically significant difference across both groups. The results indicated that students' morphological accuracy did not improve from the pre-test and post-test.

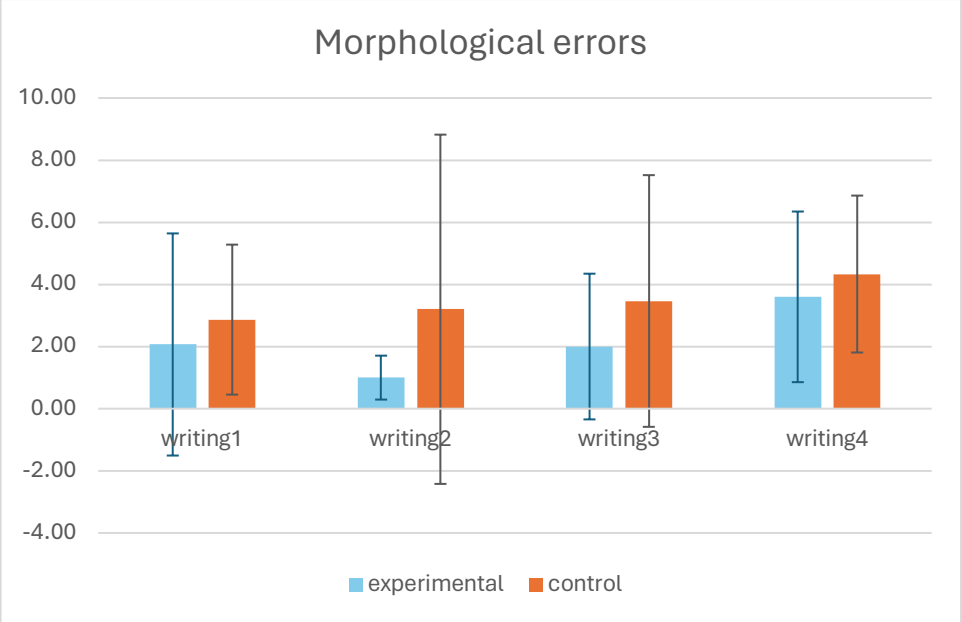


Figure 15: Bar graph showing mean morphological errors for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

As for the lexical errors, the experimental group and the control group numerically demonstrated a reduced mean lexical error from 5.13 in the pre-test to 3.13, from 7.93 to 2.07 in the post test, respectively, as shown in Figure 16. The results of the MWU test (U=148.5, z=1.499, p=.137) indicated that the differences were not statistically significant. This suggested that the intervention did not lead to a significant comparative reduction in lexical errors between the two groups. However, the outcomes of the WSR test (W=49, z=-3.513, p<.001) revealed a statistically significant reduction in lexical errors across both groups and the gains scores differed significantly from zero, indicating evidence of an overall difference in both groups, despite the lack of a significant difference (p=.137) between the groups.

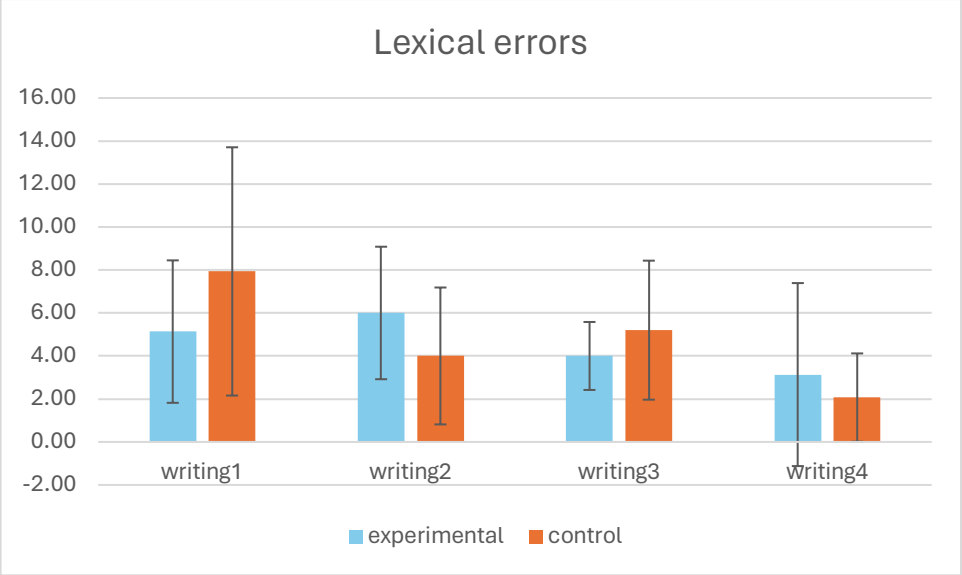


Figure 16: Bar graph showing mean lexical errors for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

In summary, as for the measurement of writing accuracy of both groups, the MWU test revealed no statistically significant differences in the change in the ratio of error-free clauses, syntactic, morphological and lexical errors between the groups. However, the WSR test, in terms of morphological and lexical errors from pre-test to post-test, indicated evidence of an overall difference in both groups.

6.4 Fluency

In terms of fluency, Figure 17 below showed that the means and standard deviation of both experimental group's and control group's writing. Numerically speaking, Figure 17 demonstrated that the mean fluency increased in the experimental group from the pre-test mean of 6.03 to the post-test mean of 6.48 and decreased in the control group from the pre-test mean of 5.82 to the post-test mean of 5.46. Regarding statistical analysis, the MWU test found that the differences in fluency gains between the experimental and control groups were not statistically significant (U=121, z=.353, p=.744). Also, the WSR test results (W=267, z=1.07, p=.284) also found that the changes in fluency in the conditions did not differ significantly from zero, indicating no evidence of an overall difference across both groups. The results suggested that there was no evidence that the groups differed in the amount they improved, nor was there evidence that, when looking across both groups, they improved from the pre-test to the post-test.

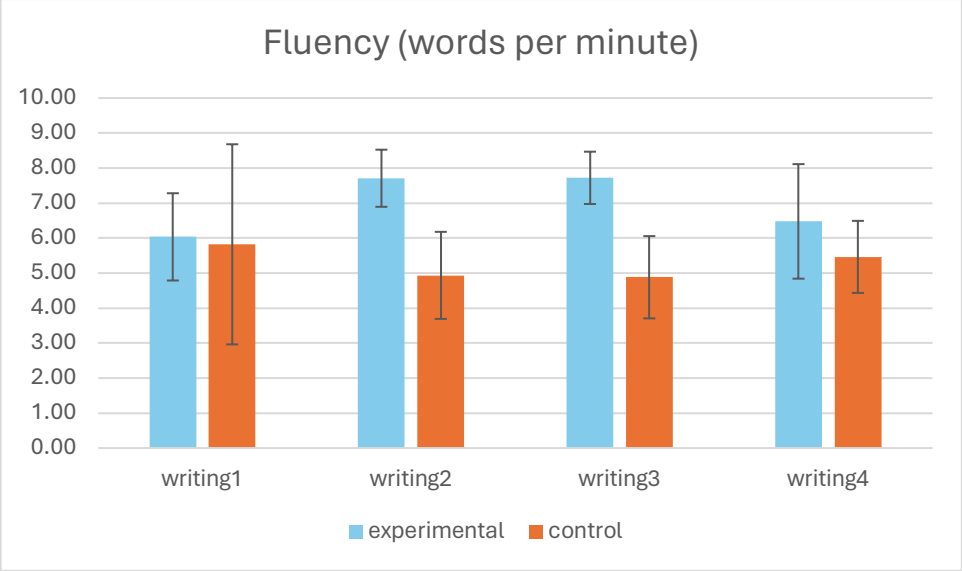


Figure 17: Bar graph showing mean fluency for the experimental and control groups in each of the four writing tasks. Error bars showing standard deviation.

6.5 Questionnaire Responses

The two questionnaires (SLWS & SLWSS) (Appendices B and C) were administered to both groups before and after the intervention. The two questionnaires utilized Likert-scale items to quantify the students' responses. To determine if there was a statistically significant difference between the pre-test and post-test scores, non-parametric tests – MWU and WSR tests – were conducted using SPSS to compare the gain scores in terms of self-efficacy and self-regulation and to investigate if there was evidence of an overall difference from the pre-test to the post-test for both groups. The data were collected from both groups and the following figures show the mean score and standard deviation of L2 writing self-regulation and self-efficacy from the pre-test and the post-test.

Concerning self-regulation, the mean scores increased numerically in both groups; the mean scores of the experimental group increased from 32.13 in the pre-test to 32.87 in the post-test and the mean scores of the control group increased from 30.47 to 30.80. Given that, the MWU test compared the gains in the scores of L2 writing self-regulation from the pre- to the post-test for the experimental and control groups and revealed no significant difference between the groups (U=138.5, z=1.106, p=.285). The WSR test found that the gains scores across both conditions significantly differed from zero (W=228.5, z = 3.362, p<.001), indicating evidence of an overall difference from the pre-test to the post-test across the conditions.

Figure 18: Bar graph showing the mean scores of L2 writing self-regulation for the experimental and control groups, as measured by the SLWS questionnaires before and after the intervention. Error bars represent the standard deviation.

Both the experimental and control groups, regarding L2 writing self-efficacy, numerically demonstrated an increase in the mean scores of L2 writing self-efficacy from the pre-test to the post-test. The experimental group had the mean scores 23.27 in the pre-test and 25.2 in the post-test and the control group numerically showed mean scores 25 in the pre-test and 27 in the post-test. The MWU test revealed that the difference was not statistically significant (U=90.0, z=-1.153, p=.367). Furthermore, the WSR test revealed that the gains scores across both conditions significantly differed from zero (W=406.0, z = 4.874, p<.001), indicating evidence of an overall difference from pre-test to post-test.

Figure 19: Bar graph showing the mean scores of L2 writing self-efficacy for the experimental and control groups, as measured by the SLWS questionnaires before and after the intervention. Error bars represent the standard deviation.

6.6 Semi-structured interview

To investigate the participants' perceptions of CW in relation to SE, SR, and writing abilities, this study employed the microgenetic method and the administered questionnaires. This approach aimed to further explain and refine our previously mentioned findings concerning CAF. The study identified themes and categories that revealed both positive and negative perceptions of the CW intervention (Tables 5 and 6). These themes and categories were used to indicate both positive and negative perceptions towards the CW intervention.

Table 5 presents the themes regarding participants' positive perceptions of the CW interview. These themes were organized into various categories. As illustrated below, five categories (self-efficacy, strong motivation, detailed attention, self-regulation, and writing efficiency) encompassing 17 themes were identified, showing the participants' positive views towards CW.

Positive student perceptions				
Self-efficacy	**Strong motivation**	**Detailed attention**	**Self-regulation**	**Writing efficiency**
Declining anxiety	Unique experience	Online cooperation	Self-monitoring	Rapid writing performance

Growing confidence	Convenient place	Collaborative macro-structure noticing	Self- and peer-correction	Further Practice
Concerted effort	Group engagement	Collaborative micro-structure noticing	Self-evaluation	
Supportive collaboration	Immediate peer-feedback		Self-managed thoughts, emotions and actions	

Table 5: Positive student perceptions towards CW (adapted and modified Rahimi & Fathi, 2022)

One category of themes addressed the positive impact of the CW intervention on the participants' increased writing self-efficacy. For instance, student A believed that the CW intervention reduced her writing anxiety:

> Interviewer: Did the collaborative writing lessons help you experience less negative feelings, such as anxiety, lack of self-confidence, fear of evaluation, etc., when you wrote in English?
>
> Student A: Before the collaborative writing lessons, I used to feel quite anxious when writing in English. You know my English is just so so. I was always worried about making mistakes and was afraid that my classmates did not like my ideas. But umm…after working with my classmates, I think it really helped me a lot. We shared ideas and gave each other feedback, which made the whole process less scary. I know we wrote it together so…yes, it reduced my fear of writing.

Another student also mentioned that writing collaboratively proffered a friendly-learning environment for them to support their peers and write collaboratively. The collaborative boosted his writing confidence.

> Interviewer: Did this support from your peers affect your writing confidence?
>
> Student C: Definitely. I think writing together boosted my confidence a lot. When we worked on the topic about …social media influencers, umm… I mean which one is more trustworthy, social media influencers, or traditional media advertisements? At first,

I couldn't think of anything! But then my groupmates were happy to share their ideas with me and then I realized that I could contribute my valuable ideas and improve my writing with their feedback. It made me believe in my own abilities more.

Another group of themes revolved around the participants' self-regulation, strong motivation, and detailed attention. Student C revealed that this CW intervention prompted him to pay more attention to self- and peer-correction, self-evaluation, and self-monitoring.

Interviewer: First, how do you plan, monitor, and evaluate your writing when working collaboratively?

Student C: When we worked on the DSE topic, we started by planning together. We brainstorm the main points together, like the credibility of influencers versus traditional ads. Because I am the team leader, we divided the tasks based on our strengths and roles, so I focused on researching the famous influencers and my two other groupmates looked into traditional media.

It was our first time to really write online and finish the writing task together… not really like doing the DSE pastpaper every day. We used the shared Google documents to monitor our progress and leave comments for each other. After drafting, we reviewed each other's work for both the overall structure and the smaller details, like grammar and clarity…and also checked if the findings and arguments correct or not.

The student C then mentioned that the CW intervention made him become more self-regulated and the group engagement also motivated him to accomplish the task.

Interviewer: That's great! Speaking of self-regulation, how do you manage your writing tasks and employ any strategies to enhance your writing?

Student C: Ar…normally, umm… I will be thinking of the objective first by analyzing the…the topic requirement with my classmates. Like this time, our goal was to argue that social media influencers are more trustworthy, we set specific targets for our arguments and evidence like finding some examples…like reports, news. Any strategies…ar…I think looking at the comments…or the tracking function is quite useful because my groupmates' feedback was quite useful. When seeing the feedback

or any suggestions on my grammar and vocab, I feel that my groupmates really worked as a team and their devotion made me feel that I should contribute more.

Lastly, concerning the theme of writing efficiency, only student B reported that the CW activity related to the theme of writing efficiency. He considered CW a unique activity, and it was conducive to his writing efficiency.

Interviewer: How different was the collaborative writing lessons from normal English writing lessons courses you had experienced in other classes?

Student B: The collaborative writing lessons were quite different from what we have. In normal English writing lessons, we are usually assigned some public exam questions and then work on our own, but sometimes, I get stuck on an idea or can't find the right words… right vocabulary to express my ideas… umm… you don't want to write if you can't express your ideas, quite frustrating… um…time-consuming, I guess? But now, we tried to write together twice, and we work as a team, which speeds up the process, because we share the ideas.

Interview: What's the difference between writing on paper and writing on Google. Document?

Student B: Writing on Google Docs is quite different from writing on paper. I think writing on it making writing much easier...I mean it is easier for us to cooperate…ar …to write together. We can all work on the same document simultaneously… much faster…after all, we have three brains, haha.

Overall, writing collaboratively via Google Docs had a positive impact on students, enhancing their motivation, efficiency in writing, attention to detail, L2 writing self-efficacy, and self-regulation. When asked if collaborative writing lessons should be incorporated into the school-based writing curriculum (interview question 8), all four interviewees expressed a unanimous preference for including collaborative writing in their regular lessons. They cited the benefits previously mentioned and the valuable new learning experiences they had gained.

However, some interviewees voiced some negative opinions about the CW intervention. Table 6 summarized the negative student perceptions towards CW. The four related categories "group heterogeneity", "different expectation", "inactive engagement" and "discouraged group members" were clustered into various themes and were shown below.

Regarding the negative student perceptions, Student C revealed that CW might not be conducive to students with a high proficient level like him because of the partial responsibility, insignificant corrective feedback and incompatible ideas. He also reckoned that some weaker students might feel more anxious and less self-confident when they worked with higher proficient students, like him.

> Student C: Sometimes, not everyone contributes equally, which can be frustrating…like in my group, I am the team leader, and one of my groupmates was quite weak in English and I guess she was quite afraid of sharing her ideas? For example, when we were debating whether social media stifles public debate, she didn't contribute much although I knew that she was doing some research online. In the whole lesson, only me and the other groupmate shared the ideas. But it doesn't mean I don't like her.
>
> Interviewer: Can you tell me more whether this weaker student provided any feedback on ideas and language using the comment and tracking functions, even though she was weaker in English?
>
> Student C: She did. Umm…sometimes I think her comments were not that useful…let me think…for example, I still remember that she suggested using "kill" to replace the word "stifle", and "offer" to replace "provide"…well… I know our English teacher always asked us to use synonyms to make our writing better, but I just think the changes were not necessary… I want to receive more comments from my teacher… not just replacing the synonym. Umm… my another groupmate and I disagreed with her ideas…and I think that these platforms actually stifle debate due to issues like echo

chambers and misinformation…but she disagreed and kept asking us…and we spent…I think…at least 5 mins explaining to her…less productive? Though I said so, I do appreciate my teammates' efforts.

Then, Student C was also afraid that his words might make his groupmate less confident in writing in English:

Student C: Umm… basically after we explained our ideas to her, she remained silent for a long time… I guess she might feel anxious or less confident in writing together.

Another student D also showed his negative perceptions towards the CW activity, in terms of group heterogeneity. He reckoned that if the groupmates were not willing to partake in the discussion and writing process, it can be detrimental to his self-confidence, and it will become more difficult to fulfill the task requirement.

Interviewer: How does the feedback from your peers during the collaborative writing process impact your self-confidence as an L2 writer?

Student D: To be honest, my two groupmates were writing very slowly…originally, we were responsible for writing different paragraphs…and we should give feedback to each other. But when I finished my part, they were still writing… and it caused some delay in our writing…Umm…I think they were not devoted to writing and working with them somehow makes me feel less confident and less willing to write…like the feedback was not constructive or too simple. For example, my groupmates simply commented on my ideas with "yes, agree", without any explanation …. suggestions.

To summarize, while CW has some potential benefits, Students C's and D's experiences revealed some drawbacks, especially when group members have varying levels of proficiency and are less committed to writing collaboratively. The drawbacks they experienced included partial responsibility, ineffective feedback, and decreased self-confidence and motivation.

CHAPTER FIVE
DISCUSSION AND CONCLUSION

7. Discussion

This chapter discusses the results of this study in relation to the research questions and previous studies on collaborative writing. It is important to note that the results concerning CAF were based on writing 1 and 4 written individually by both groups while the microgenetic method was used to analyze writing 2 and 3 written collaboratively by the experimental group only. The first research question explores how a Google Docs-supported collaborative writing approach, emphasizing peer feedback and revision, affects the subsequent individual writing performance of Band 1 Secondary Four students in Hong Kong. To address this, the CAF measurement was employed on the individual essays written before and after the intervention, with empirical evidence and detailed discussions provided in Section 6.1. The second research question, "How does Google Docs-mediated collaborative writing influence the development of writing self-regulation among Band 1 Secondary Four students in Hong Kong?" and the third question, which is about the impacts of Google Docs-mediated CW on Band 1 Secondary Four students' L2 writing self-efficacy, are addressed in Sections 6.2 and 6.3, respectively. These sections analyze the empirical data from the two questionnaires and qualitative findings from the interviews.

7.1 Effects of CW on complexity, accuracy and fluency

To investigate the impacts of CW on the writing performance of Band 1 secondary four students, syntactic complexity was measured. First, concerning the test results of the number of T-units, the number of clauses, and the number of words per clause, this study compared the gains between the two groups and found that there was no statistically significant difference in improvements concerning the above three items between the experimental and control groups. It also assessed the change from pre-test to post-test across both conditions and found that no statistically differences were found in the number of T-units, the number of clauses and the number of words per clause, suggesting that the intervention did not effectively enhance syntactic complexity as measured by these three items. Interestingly, in terms of the gains in the ratio of clauses per T-units, this study revealed a statistically significant difference between the two groups. However, when examining the pre-test and post-test scores of each group using the WSR test, the result was mixed. In the experimental group, no significant difference was found between the pre-test and post-test scores, indicating that the intervention did not lead to

improvement in the experimental group; conversely, the control group showed a statistically significant difference between the pre-test and post-test scores and therefore it suggests that the control group enhanced the ratio of clauses per T-units in their writing without the intervention. There is no obvious explanation for this pattern, which could potentially be attributed to a Type 1 error.

To shed light on these findings and explore potential benefits of the Google Docs-mediated CW intervention that may not have been captured by the quantitative measures, this study turns to the qualitative data obtained through the microgenetic method. During the collaborative writing process, the experimental group initiated 14 mediations related to changes in the structure of their writing, focusing on aspects such as active/passive voice (3 mediations), relative clauses (6 mediations), and word order (5 mediations). These mediations suggest that the intervention encouraged the students to actively engage with and discuss various elements of syntactic complexity, even if these discussions did not directly translate into measurable improvements in the post-test.

In terms of lexical complexity, although a slight numerical increase was observed in the New General Service List and non-General Service List coverage from the pre-test to the post-test, these changes were not statistically significant. Overall participants improved but there was no evidence that one group improved more than the other. Despite the lack of significant quantitative results, the qualitative data derived from the microgenetic analysis, and the interview provide some interesting insights into the impacts of the intervention on the experimental group's lexical awareness and linguistic behaviour. For example, the microgenetic analysis showed a notable shift in the experimental group's attention towards word choice and the use of vocabulary to enhance accuracy and variety. Out of 162 language-related mediations in writing 2 and 3, nearly half (83 instances) focused on vocabulary enhancement like using synonym or more formal words. This indicates a conscious effort by the students to refine their lexical choices and enhance formality and precision of the writing. The interview further corroborated these findings as one of the interviewees revealed that he valued and benefited from the peer feedback on the use of vocabulary and grammar, and such interaction initiated by the intervention encouraged him and his groupmates to use different vocabulary more accurately, in terms of formality and variety.

As for morphological complexity (nouns), the study found that there was a statistically improvement in MC (nouns) in both the experimental and control groups over the course of this study. Overall participants improved but there was no evidence that one group improved more than the other. Given the statistical results, this study cannot conclusively state that the CW intervention enhances MC (nouns) in the students' writing because the improvement in the control group suggests that there are other factors such as topic familiarity, natural development in writing skills over time, other than the intervention, may have contributed to the improvement.

Regarding accuracy, the differences in the changes between the two groups were not statistically significant for any of the accuracy measures, suggesting that the CW intervention did not have a significant impact on improving the accuracy of the experimental group, compared to the control group. A statistically significant reduction in lexical errors across both groups was found, suggesting that the intervention may have had positive impacts on the experimental group's lexical accuracy. It is important to note that the improvement was not significantly different between the experimental and control groups, and it is inconclusive to say that the intervention was effective on the students' accuracy as both groups improved. Surprisingly, this study indicated a statistically significant difference, but in the opposite direction of improvement, with both groups showing an increase in morphological errors. Despite the lack of statistically significant improvements in most accuracy measures, the qualitative data obtained through the microgenetic analysis sheds light on the potential benefits of the intervention on the students' L2 accuracy. To be specific, the analysis of language-related mediations in writings 2 and 3 revealed that the majority of mediations (83 out of 162) were related to word choice. Moreover, the high number of mediations related to word choice suggests that the interaction in the CW intervention encouraged the students to engage in the revision process and discussions about the use of vocabulary. However, since the control group also demonstrated the same reduction in lexical errors after the intervention, the effectiveness of the Google Docs-mediated CW on the students' accuracy remains uncertain.

Lastly, this study also found that the CW intervention did not improve students' writing fluency. The study found no evidence of an overall improvement across both groups, thus indicating that the intervention did not effectively enhance the students' ability to write more words per minute. Despite the quantitative data, the interview provides some insights into the benefits of the intervention. The students reported that they could write faster when engaged in the CW activity. They attributed this improvement in fluency to the opportunity to share ideas

with their peers and seek assistance from more capable group members when facing language-related challenges, such as a lack of vocabulary and writer's block. Sharing ideas and receiving input from others may contribute to an increased writing fluency. While the quantitative results did not demonstrate statistically significant improvements in writing fluency, the qualitative data from the interview sheds light on the positive impact of the intervention on their writing fluency.

To answer the research question 1: "How does a Google Docs-supported collaborative writing approach, with a specific focus on peer feedback and revision, impact the subsequent individual writing performance of Band 1 secondary four students in Hong Kong?", the findings suggest that while the intervention did not result in statistically significant improvements in most of the quantitative measures of CAF, the microgenetic analysis in writing 2 and 3 revealed that the students actively engaged in discussions and initiated mediations related to content, language and organization. The interview further highlighted the potential benefits of the intervention, including increased writing fluency, enhanced lexical awareness, and improved ability to generate ideas and overcome language-related challenges through peer collaboration. Although the findings from this study appear beneficial, they do not conclusively demonstrate improvement in the participants' individual writing abilities. This limitation can be attributed to a few factors, including the need for modelling collaborative interactions, providing sufficient practice sessions, and extending the duration of implementing collaborative writing (CW), as pinpointed by Storch (2013). In contrast, the positive and statistically significant results observed in Rahimi's & Fathi's (2022) research may also be attributed to similar factors. To be specific, participants in their study received detailed instructions about wiki-mediated CW and the instructional design before the intervention. Furthermore, they received 12 ninety-min sessions of CW lessons were allowed to collaboratively write, edit, and discuss ideas outside the classroom, thus extending their exposure time to collaborative writing practices and allowing more opportunities to model or adjust their collaborative interactions. Conversely, in this study, the students were allocated only 70 minutes for all writing tasks and were restricted from writing outside the classroom. As a result, they did not receive sufficient training sessions on Google Docs-mediated CW, nor did they have sufficient time to model interactions with their groupmates. Engaging in collaborative writing might potentially enhance individual writing capabilities, but the mere two sessions conducted in this study proved insufficient for a measurable impact. Therefore, to more accurately assess changes in CAF in their writing, this study recommends repeating the

CW activity with a longer intervention stage, involving students in CW repeatedly over a few months rather than just twice. Taking the exploratory nature and the sample size of this study into consideration, the speculations remain tentative and warrant further investigation, which align with some previous studies reporting no language learning advantage for group over individual task completion when comparing the pre-test and (delayed) post-test performance (Kuiken & Vedder, 2002; Nassaji & Tian, 2010; Reinders, 2009).

7.2 Impacts of CW on self-regulation

In terms of the positive impacts of technology-mediated CW on students' self-regulation, previous studies (Boykin et al., 2019; Su et al., 2018) found that online writing tools like Wiki can facilitate students' self-regulation and provide a convenient and user-friendly space for students to apply self-regulatory strategies including setting goals, brainstorming, organizing ideas, and engaging in self-monitoring and self-evaluation. (Rahimi & Fathi, 2022). No significant difference was found between the groups in terms of the gains in self-regulation, suggesting that the intervention did not have a significantly greater impact on the experimental group's self-regulation compared to the control group. In fact, overall both groups improvement in their SR scores from pre- to post test, suggesting that some factor other than engaging in the collaborative writing intervention was leading to this benefit. Moreover, the improvement in their SR scores can be attributed to an effect of retesting. The students might feel that they needed to answer what they were supposed to give as the researcher was also their English teacher.

On the other hand, the qualitative data from the semi-structured interview might also shed light on the students' enhanced self-regulation and how the intervention influenced their self-regulatory abilities. For instance, Student C reported that using shared Google documents facilitated the setting of writing goals, monitoring progress, and providing peer feedback, which enhanced his self-regulation skills. Furthermore, other students also noted that Google Docs enabled more proficient students to employ self-regulatory writing strategies to assist their less capable peers. These strategies included setting writing goals, generating and organizing ideas, and evaluating text quality during collaboration with lower-proficiency groupmates. Moreover, the microgenetic analysis also revealed that the mediations concerning content, language, and organization may have enabled students to more actively control, manage, and plan their writing tasks, thereby improving their self-regulatory capabilities (Rahimi & Fathi, 2022), because they needed to evaluate their groupmates' writing before

initiating the mediations, which required them to engage in self-regulatory processes such as self-monitoring and self-evaluating. This research cannot conclusively affirm that the effectiveness of CW on L2 writing SR is beneficial. The findings in this study contrast with the research design and outcomes of previous studies (Qiu & Lee, 2020; Rahimi & Fathi 2022). For instance, Qiu & Lee (2020) required the participants to write reflective notes after each session to express their feelings and opinions about peer CW. They were also allowed to use their first language, Chinese, during discussions and in reflective notes. Furthermore guiding questions were proffered for the participants to help facilitate the drafting of reflections, making it easier for them to self-regulate their learning processes. In contrast, the methodology of this study required the participants to complete the questionnaire before the intervention (writing 1) and after the intervention (writing 3). There is a possibility that they might have felt inclined to report higher scores on the Likert scale when completing the same questionnaire repeatedly. Unlike in Qiu & Lee's study (2020), they could not refer to their reflective notes while responding to the questionnaires. This difference in approach may have impacted their ability to self-regulate effectively during the study. Therefore, to answer the research question two, the effectiveness of the intervention on the students' L2 SR appears to be mixed, as evidenced by the quantitative and qualitative data, and therefore, no definite conclusion can be drawn.

7.3 Impacts of CW on self-efficacy

As for the final research question, no statistically significant difference was found in SE scores between the groups, suggesting that the CW intervention did not have a significantly greater impact on the experimental group's self-efficacy compared to the control group. Although overall the participants improved, there was no evidence that one group improved more than the other. Therefore, it is uncertain to assert that this intervention was effective. To further investigate the impact of the intervention on their self-efficacy, it is of paramount importance to consider the qualitative data obtained in the interview. For instance, the students reported reduced anxiety and increased confidence, when they cooperated with their groupmates as their groupmates provided the immediate, supportive feedback through the comment and tracking function. These narratives and learning experiences underscore not only an enhancement in their L2 writing SE but also an enjoyment for the CW intervention. The effectiveness of Google Docs-mediated CW on the students' L2 SE can be influenced by factors such as the use of the first language (Fung, 2010) and the method of grouping (Anderson, 2019). For example, in the study by Qiu & Lee (2020), the students were permitted to use their L1 during the intervention, whereas the use of L1 was prohibited in the current research. The

interviews from this study revealed that the students felt more anxious and less confident in participating in the CW activity when faced with language barriers due to their limited L2 proficiency. Additionally, the composition of groups can significantly affect L2 SE, particularly for less proficient students who are grouped with more proficient peers. Such grouping can alter interaction patterns, potentially inducing anxiety and diminishing L2 SE in some students. In this research, the prohibition of L1 use and the grouping decisions made by the teacher and researcher may have influenced the outcomes. Given the mixed findings from the quantitative and qualitative data, the effectiveness of the Google Docs-mediated CW intervention on the band 1 secondary four students' L2 SE remains inconclusive.

8. Pedagogical implications

Although the quantitative results may not conclusively prove the effectiveness of the intervention on the students' writing performance, L2 SR and SE, the students in the interviews expressed their positive attitudes towards the intervention. As such, it is useful to explore the pedagogical implications of the intervention and consider how educators and teachers can efficaciously integrate this approach into the curriculum and support the holistic development of students' writing skills, L2 SR and SE.

8.1 Patterns of interaction

In the technology-mediated CW, one of the problems faced by students is the patterns of interaction when they work in groups, especially their contribution and engagement with each other's contributions (Storch, 2013). In this study, the students in the interview reported similar experiences that their groupmates remained silent after their suggestions were not accepted. It is important to note that simply assigning learners to work together in small groups on writing tasks does not necessarily mean that learners will collaborate with one another to write online, and therefore it is necessary to encourage the formation of collaborative relationships. As such, modelling collaborative interaction, providing practice sessions and monitoring the nature of learners' interactions are suggested (Storch, 2013). To be specific, first, teachers can use a short video or the researcher/teacher modelling collaborative dialogues as a form of modelling. Previous studies have shown that the modelling can encourage learners to adopt a collaborative pattern of interaction in L2 contexts, with adult learners (Kim & McDonough, 2011), and in L1 contexts with young learners (Dale, 1994).

Moreover, assigning students with different English levels to work in groups does not guarantee that they will know how to handle the group dynamics and collaborate. Therefore, training and practice are necessary to develop skills such as teamwork, task management and how to negotiate roles and responsibilities (Dovey, 2006). Specifically, pre-task modelling of collaboration and training on providing peer feedback (Arnold et al., 2009) should be proffered for students before the intervention.

In order to ensure the success of CW, teachers should assume the pivotal roles of a monitor and facilitator. While some learners may cooperate effectively with their peers, others may lack motivation and tend to become free riders or remain silent, as reported by the student in the interview. To prevent this, teachers should actively supervise and provide guidelines on interactions and feedback. As a facilitator, teachers can encourage them to contribute ideas, offer help, provide suggestions, and seek clarification from their peers when needed, thereby fostering a two-way interaction that characterizes successful collaboration (McInnerney & Roberts, 2009). Kessler (2009) also recommends that teachers should be involved in the activity to prompt students to participate in group online discussions themselves. This careful monitoring can help address potential problems related to group dynamics early in the intervention. For instance, one student in this study reported that their groupmates remained passive throughout the discussion, a situation that could have been mitigated by teacher monitoring and intervention. As Fullan (2009) and Farrell and Jacobs (2016) emphasize, the successful implementation of cooperative learning depends heavily on how teachers facilitate effective interaction among students. Therefore, to maximize the benefits of CW interventions, teachers should actively engage in monitoring, facilitating, and guiding student collaboration.

8.2 The use of L1

In this study, the school is an EMI school, and therefore the use of L1 in English lessons is not allowed. As discussed in Section 6.3, the use of L1 can influence learners' SE. During the interview, the students reported experiencing language barriers and a lack of vocabulary, which negatively impacted their L2 SE. Moreover, as CW requires speaking skills, and speaking is known to be an anxiety-provoking activity in L2 English classrooms (Horwitz et al., 1986), the low-achieving students may feel anxious about expressing ideas in English during peer interactions. Consequently, this challenge may negatively affect their participation in the lesson when English is the sole medium of instruction and communication in the EMI school in Hong Kong. To avoid the challenge of speaking English, some learners, particularly

those with low proficiency, prefer to use their L1 to communicate with peers during the CW activity. This phenomenon has been observed by some researchers (Le & McDonald, 2004; Platt & Brooks, 2002). In view of the potential benefits of the use of L1, Fung (2010) suggests that the use of L1 should not be entirely prohibited, as it can assist weaker learners in the learning process by facilitating idea generation and effectively conveying the meaning of unknown words in L2 through translation.

8.3 Grouping in CW

Teachers may face challenges in grouping learners when implementing CW in an English classroom, especially a group of students with mixed abilities. It is inevitable that students perform differently due to their English levels and individual differences (Anderson, 2019). For instance, low-achieving students may become free riders because they are not able to proofread the work or contribute ideas and therefore higher achievers may bear a greater workload and responsibility in contributing ideas and proofreading the written text. In this study, a similar situation has also been reported by the experimental group. Therefore, this study, making reference to Vygotsky's concept of the ZPD, suggests that teachers should appoint high achievers as leaders to help low achievers in each group and provide guidelines and instructions on how students with different abilities should work as a team. The ZPD represents the difference between what a learner can do without help and what they can do with guidance from a more skilled learner (Vygotsky, 1978). In CW, teachers should keep in mind that students learn from their "more capable peers" (Nordlof, 2014). By grouping high achievers with low achievers, teachers can create opportunities for peer-assisted learning within the ZPD.

9. Limitations of the Present Study

9.1 Sample size

Firstly, the participant sample size for this study was limited to 30 Band 1 secondary four students in Hong Kong who responded to the questionnaires. In the experimental group of 15 students, only four students were interviewed. Regarding the writing samples, the experimental group produced 30 individual samples and 10 collaboratively written samples, while the control group contributed 60 individual samples for analysis. Specifically, only 10 collaboratively written samples (writing 2 and 3) were analyzed using the microgenetic method. Therefore, the findings of this study are contextually limited to the interpretations of the 30 participating students and may not necessarily be generalizable to the broader population of

Band 1 secondary four students in Hong Kong or in other contexts due to the limited number of essays for doing statistical analyses comparing scores on the writing 2 and 3 essays. Similar limitations have been encountered in other studies (Abadikhah, 2012; Fortune, 2005; Wang, 2015).

9.2 The measurement tools – LancsLex & Morpho Complexity Tool

In terms of the instrument, although LancsLex can assess lexical complexity using NGSL, there are some limitations. First, it only considers the frequency of words and does not account for other factors that contribute to lexical complexity, such as word length, morphological complexity, or semantic complexity. Second, the NGSL is based on a specific corpus of texts and may not fully represent the vocabulary used in all contexts or genres. Lastly, this study also found that the spelling mistakes made by the participants were also considered as the non-NGSL vocabulary, resulting in slight inaccuracy when measuring lexical complexity. This issue arises because LancsLex analyzes the writing based on exact words in its database. When a spelling error occurs, the misspelled word does not match a word in the NGSL, leading LancsLex to catergorize the misspelled word as a non-NGSL word. This misclassification has a direct impact on lexical complexity of a writing. For instance, if a student's writing contains many spelling mistakes, these mistakes are counted as a use of more complex vocabulary, thus resulting in higher lexical complexity.

Additionally, the Morpho Complexity Tool used for assessing morphological complexity exhibited its own constraints. The primary limitations include the lack of systematic implementation in the analysis of nouns and periphrastic morphemes, and the preliminary nature of its analysis for the English language (Pallotti & Brezina, 2023).

9.3 Insufficient duration of the CW intervention for measuring L2 writing performance, SE and SR

In addition, the duration of the intervention of CW or technology-mediated CW only lasted a few weeks in the previous studies (Abadikhah, 2012; Dobao, 2012; Fortune, 2005). In this study, similarly, the technology-mediated CW only took place in one academic term (2 months). This duration may be considered insufficient to thoroughly measure the longitudinal improvement in the students' writing performance, L2 SR and SE (Rahimi & Fathi, 2022) as longitudinal studies allow researchers or teachers to monitor the evolution of interaction patterns and group dynamics, and the changes in their SR and SE over time.

9.4 Limitations of the semi-structured interview and questionnaires

First, because the interviewer has dual roles as the researcher and English teacher of both the experimental and control groups, researcher bias is easily produced. The role as a teacher can influence how the students answered the questions and how the responses were interpreted. Therefore, it may potentially lead to biased data collection, especially in the interviews. The participants might want to alter their responses in the interview based on the relationship with the teacher or researcher.

Another limitation that might affect the findings of the study is the use of L2. The questionnaires were written in English language, not in their mother tongue, Cantonese, and therefore there was a risk that the students might not fully understand the questions. This can result in inaccurate responses. Moreover, the interview was also conducted in their L2 as the school was an EMI school, which required all the students to use English in the English lessons and English-related activities. The use of L2 might cause anxiety, or misinterpretations to the students when they were filling in the questionnaires and having the interview.

9.5 Task type

Previous studies conducted by Lee (2010), Lund (2008), and Mak & Coniam (2008) contend that the nature of the task significantly influences the level of engagement and collaboration in the technology-mediated CW learning environment, rather than the technology itself. As discussed previously, tasks that are focused on meaning, such as jigsaw puzzles and data commentary texts, tend to elicit fewer LREs compared to those that emphasize language and grammar, such as dictogloss, editing, and cloze exercises. Furthermore, these meaning-focused tasks are less likely to generate grammatical LREs than lexical LREs. However, it is also noted that meaning-focused tasks enhance learners' awareness of form-meaning connections, thereby facilitating a deeper engagement with language choices (Storch, 2013). In this study, the four writing tasks were meaning-focused and the data collected from questionnaires and interviews suggest that the intervention may have had a positive impact on students' attitudes towards CW, L2 writing SE and SR. It is likely that the outcomes of the intervention could be generalizable, and the results might have differed if the tasks had been grammar-focused. For example, grammar-focused tasks emphasize linguistic accuracy over fluency, and it is conceivable that different aspects of writing proficiency, such as syntactic and morphological accuracy, would have been more significantly impacted if this research had adopted grammar-focused tasks instead of meaning-focused ones. Furthermore, this research

utilized the microgenetic method for qualitative analysis; the number of language-related mediations were fewer than those typically observed in grammar-focused tasks. As a result, the task type in this study may limit the measurement of linguistic accuracy.

10. Conclusion

To conclude, while technology-mediated CW has been advocated by numerous studies for its beneficial impacts on learners' CAF, L2 SR and SE, this study did not find evidence of the benefits for Band 1 Secondary Four students. This study does not align with most of the studies on CW (Fortune, 2015; Dobao, 2012; Qiu & Lee, 2020) that reported positive statistically significant outcomes on CAF. However, some positive learning experiences were reported by the students, and they showed positive attitudes towards CW concerning their writing performance, L2 SR and SE in the interviews.

In view of the outcomes and the limitations of this study, for future research, this study would like to provide a few suggestions. First, it would be prudent to include students from Band 2 and Band 3 to prevent overgeneralization from a sample consisting solely of high-achieving Band 1 students. Second, the sample size should be expanded to enhance the reliability of the results and avoid generalization (Li & Zhang, 2021). Third, while this study focused on the meaning-focused genre of argumentative writing, future research should explore other meaning-focused genres and include grammar-focused tasks. Fourth, considering that this research was conducted over a span of only two months, implementing longitudinal studies would provide deeper insights into the long-term effects of the CW intervention. Lastly, concerning the linguistic environment, the current study was conducted in a local band 1 school where the use of Cantonese was restricted in the lessons. Future studies should consider the implications of allowing or restricting L1 use, or implementing bilingual approaches during the CW implementation. This could include examining how L1 inclusion affects learners' interaction patterns, L2 SR and SE, and overall learning outcomes.

References

Abadikhah, Shirin. (2012). The effect of mechanical and meaningful production of output on learning English relative clauses. *System (Linköping)*, *40*(1), 129-143.

Alberth. (2019). Use of Facebook, students' intrinsic motivation to study writing, writing self-efficacy and writing performance. *Technology, Pedagogy and Education*, *28*(1), 21–36. https://doi.org/10.1080/1475939X.2018.1552892

Allan, D. (2004). *Oxford placement test*. Oxford: Oxford University Press.

Anderson, J. (2019). Cooperative learning: principles and practice. *English Teaching Professional*, *121*, 4-6.

Andujar, A. (2016). Benefits of mobile instant messaging to develop ESL writing. *System (Linköping)*, *62*, 63–76. https://doi.org/10.1016/j.system.2016.07.004

Arnold, N., & Ducate, L. (2019). *Engaging language learners through CALL : From theory and research to informed practice* (Advances in CALL research and practice). Sheffield, UK.

Arnold, N., Ducate, L. and Kost, C. (2009) Collaborative writing in wikis: Insights from a culture project in a German class. In L. Lomicka and G. Lord (eds) *The next generation: Social networking and online collaboration in foreign language learning* (pp. 115-144). San Marcos, TX: CALICO.

Avci, H., & Adiguzel, T. (2017). A Case Study on Mobile-Blended Collaborative Learning in an English as a Foreign Language (EFL) Context. *International Review of Research in Open and Distance Learning*, *18*(7), 45–58. https://doi.org/10.19173/irrodl.v18i7.3261

Avci, Hulya, & Adiguzel, Tufan. (2017). A Case Study on Mobile-Blended Collaborative Learning in an English as a Foreign Language (EFL) Context. *International Review of Research in Open and Distance Learning*, *18*(7), 45-58.

Aydin, Z., & Yildiz, S. (2014). Using wikis to promote collaborative efl writing. *Language Learning & Technology*, *18*(1), 160–180.

Banegas, Darío Luis, Loutayf, María Soledad, Company, Susana, Alemán, María José, & Roberts, Grisel. (2020). Learning to write book reviews for publication: A collaborative action research study on student-teachers' perceptions, motivation, and self-efficacy. *System*, *95*, System. Volume 95 (2020).

Bikowski, D., & Vithanage, R. (2016). Effects of web-based collaborative writing on individual l2 writing development. *Language Learning & Technology*, *20*(1), 79–99.

Blake, Robert. (2000). Computer Mediated Communication: A Window on L2 Spanish Interlanguage. *Language Learning & Technology*, *4*(1), 120-136.

Boticki, Ivica, Looi, Chee-Kit, & Wong, Lung-Hsiang. (2011). Supporting Mobile Collaborative Activities through Scaffolded Flexible Grouping. *Educational Technology & Society*, *14*(3), 190-202.

Boykin, Andrea, Evmenova, Anya S., Regan, Kelley, & Mastropieri, Margo. (2019). The impact of a computer-based graphic organizer with embedded self-regulated learning strategies on the argumentative writing of students in inclusive cross-curricula settings. *Computers and Education*, *137*, 78-90.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, *36*(1), 1–22. https://doi.org/10.1093/applin/amt018

Brezina, V., & Gablasova, D. (2017). How to Produce Vocabulary Lists? Issues of Definition, Selection and Pedagogical Aims. A Response to Gabriele Stein. *Applied Linguistics*, *38*(5), 764–767. https://doi.org/10.1093/applin/amx022

Brezina, V., & Pallotti, G. (2015). Morphological complexity tool. http://corpora.lancs.ac.uk/vocab/analyse_morph.php

Brezina, V., & Pallotti, G. (2023). Computing the Morphological Complexity Index. http://corpora.lancs.ac.uk/vocab/analyse_morph.php

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, *105*(1), 25–38.

Camps, A., Guasch, O., Millan, M. and Ribas, T. (2000). Metalinguistic activity: The link between writing and learning to write. In A. Camps and M. Milan (eds) *Metalinguistic activity in learning to write* (pp. 103–124). Amsterdam: Amsterdam University Press.

Caplan, Nigel A., & Farling, Monica. (2017). A Dozen Heads Are Better Than One: Collaborative Writing in Genre-Based Pedagogy. *TESOL Journal*, *8*(3), 564-581.

Carver, Julie, & Kim, YouJin. (2020). French Learners' Past-Tense Development through Collaborative Writing Tasks: The Role of Procedural and Content Repetition. Canadian Modern Language Review, 76(2), 114-138.

Chang, G. and Wells, G. (1988). The literate potential of collaborative talk. In T. Maclure, T. Phillips and A. Wilkson (eds) *Oracy Matters* (pp. 95–109). Stony Stratford, England: Open University Press.

Chen, W. (2019). An exploratory study on the role of L2 collaborative writing on learners' subsequent individually composed tests. *The Asia-Pacific Education Researcher*, *28*(6), 563-573.

Cho, Moon-Heum, & Lim, Seongmi. (2017). Using regulation activities to improve undergraduate collaborative writing on wikis. *Innovations in Education and Teaching International*, *54*(1), 53-61.

Colina, Ana Alegría De La, & Mayo, María Del Pilar García. (2018). Attention to Form Across Collaborative Tasks by Low-Proficiency Learners in an EFL Setting. In *Investigating Tasks in Formal Language Learning* (pp. 91-116). Bristol, Blue Ridge Summit: Multilingual Matters.

Csizér, Kata, & Tankó, Gyula. (2017). English Majors' Self-Regulatory Control Strategy Use in Academic Writing and its Relation to L2 Motivation. *Applied Linguistics*, Amv033. discourse. *System*, *30*(2), 275-288.

Dale, H. (1994). Collaborative Writing Interactions in One Ninth-Grade Classroom. *The Journal of Educational Research (Washington, D.C.)*, *87*(6), 334–344. https://doi.org/10.1080/00220671.1994.9941264

Dobao, Ana Fernández. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, *21*(1), 40-58.

Donato, Richard. (2004). Aspects of Collaboration in Pedagogical Discourse. *Annual Review of Applied Linguistics*, *24*, 284-302.

Dooly, Melinda, & Sadler, Randall. (2016). Becoming little scientists: Technologically enhanced project-based language learning. *Language Learning & Technology*, *20*(1), 54-78.

Dovey, T. (2006). What purposes, specifically? Re-thinking purposes and specificity in the context of the 'new vocationalism.' *English for Specific Purposes (New York, N.Y.)*, *25*(4), 387–402. https://doi.org/10.1016/j.esp.2005.10.002

Ede, L., & Lunsford, A. (1990). *Singular texts/plural authors: Perspectives on collaborative writing*. Southern Illinois University Press.

Edstrom, Anne. (2015). Triads in the L2 classroom: Interaction patterns and engagement during a collaborative task. *System (Linköping)*, *52*, 26-37.

Elabdali, Rima. (2021). Are two heads really better than one? A meta-analysis of the L2 learning benefits of collaborative writing. *Journal of Second Language Writing*, *52*, 100788.

Esteban, Noelia Martinez, & De Larios, Julio Roca. (2010). The Use of Models as a Form of Written Feedback to Secondary School Pupils of English. *International Journal of English Studies*, *10*(2), 143.

Fathi, Jalil, & Rahimi, Masoud. (2022). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: A case of EFL students. *Computer Assisted Language Learning*, *35*(7), 1668-1706.

Fathi, Jalil, Mohebiniya, Sara, & Nourzadeh, Saeed. (2019). Enhancing Second Language Writing Self-regulation Through Self-assessment and Peer-assessment: A Case of Iranian EFL Learners. *International Journal of Applied Linguistics & English Literature*, *8*(3), 110.

Fernandez Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, *21*(1), 40–58. https://doi.org/10.1016/j.jslw.2011.12.002

Fortune, Alan. (2005). Learners' Use of Metalanguage in Collaborative Form-focused L2 Output Tasks. *Language Awareness*, *14*(1), 21-38.

Fung, Y. M. (2010). Collaborative Writing Features. *RELC Journal*, *41*(1), 18-30.

García Mayo, María. (2002). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research*, *37*(3), 323 341.

Guo, Sijia, & Möllering, Martina. (2016). The implementation of task-based teaching in an online Chinese class through web conferencing. *System (Linköping)*, *62*, 26-38.

Han, Jiwon, & Hiver, Phil. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing*, *40*, 44-59.

Hsu, H. C., & Lo, Y. F. (2018). Using wiki-mediated collaboration to foster L2 writing performance. *Language Learning & Technology*, *22*(3), 103–123.

Hsu, H.-C. (2019). Wiki-mediated collaboration and its association with L2 writing development: an exploratory study. *Computer Assisted Language Learning*, *32*(8), 945–967. https://doi.org/10.1080/09588221.2018.1542407

Kessler, G. (2009). Student-Initiated Attention to Form in Wiki-Based Collaborative Writing. *Language Learning & Technology*, *13*(1), 79–95.

Kessler, Greg, & Bikowski, Dawn. (2010). Developing collaborative autonomous learning abilities in computer mediated language learning: Attention to meaning among students in wiki space. *Computer Assisted Language Learning*, *23*(1), 41-58.

Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research : LTR*, *15*(2), 183–199. = https://doi.org/10.1177/1362168810388711

Kim, YouJin, & McDonough, Kim. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research: LTR*, *12*(2), 211-234.

Kim, Youjin. (2008). The Contribution of Collaborative and Individual Tasks to the Acquisition of L2 Vocabulary. *The Modern Language Journal (Boulder, Colo.)*, *92*(1), 114-130.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*, 390-403.

Krashen, S. (1981). Second language acquisition and second language learning (1st ed., Language teaching methodology series). Oxford ; New York: Pergamon Press.

Krashen, S. (1982). Principles and practice in second language acquisition (Language teaching methodology series). Oxford: Pergamon.

Krashen, S. (1985). The input hypothesis: Issues and implications. London: Longman.

Kuiken, Folkert, & Vedder, Ineke. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, *37*(3), 343-358.

Laet, T., Klemke, R., Alario-Hoyos, C., Hilliger, I., & Ortega-Arranz, A. (2021). *Technology enhanced learning for a free, safe, and sustainable World : 16th European Conference on Technology Enhanced Learning, EC-TEL 2021, Bolzano, Italy, September 20-24, 2021, Proceedings* (Lecture notes in computer science; 12884). Cham, Switzerland.

Lai, Chiu-Lin. (2021). Effects of the group-regulation promotion approach on students' individual and collaborative learning performance, perceptions of regulation and regulation behaviours in project-based tasks. *British Journal of Educational Technology*, *52*(6), 2278-2298.

Landrieu, Yana, De Smedt, Fien, Van Keer, Hilde, & De Wever, Bram. (2023). Argumentation in collaboration: The impact of explicit instruction and collaborative writing on secondary school students' argumentative writing. *Reading & Writing*, 1-28.

Lantolf, J.P. (2000). Introducing sociocultural theory. In J.P. Lantolf (ed.) *Sociocultural theory and second language learning* (pp. 1–26). Oxford: Oxford University Press.

Law, Victor, Ge, Xun, & Eseryel, Deniz. (2016). The Development of a Self-regulation in a Collaborative Context Scale. *Technology, Knowledge and Learning*, *21*(2), 243-253.

Le, P. H. H., & McDonald, G. (2004). Mediation through the first language: A sociocultural study of group work in Vietnam. *New Zealand Studies in Applied Linguistics*, *10*(1), 31-49.

Lee, L. (2002) Synchronous online exchanges: A study of modification devices on nonnative

Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning & Technology*, *8*(1), 83–100.

Lee, L. (2010). Exploring Wiki-Mediated Collaborative Writing: A Case Study in an Elementary Spanish Course. *CALICO Journal*, *27*(2), 260–276. https://doi.org/10.11139/cj.27.2.260-276

Lee, Lina. (2007). Fostering Second Language Oral Communication Through Constructivist Interaction in Desktop Videoconferencing. *Foreign Language Annals*, *40*(4), 635-649.

Lee, Lina. (2008). Focus-on-Form Through Collaborative Scaffolding in Expert-to-Novice Online Interaction. *Language Learning & Technology*, *12*(3), 53-72.

Leeser, Michael J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research: LTR*, *8*(1), 55-81.

Li, Hui Helen, & Zhang, Lawrence Jun. (2021). Effects of structured small-group student talk as collaborative prewriting discussions on Chinese university EFL students' individual writing: A quasi-experimental study. *PloS One*, *16*(5), E0251569.

Li, Mimi. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, *31*(8), 882-904.

Li, X., & Chu, S. K. W. (2018). Using design-based research methodology to develop a pedagogy for teaching and learning of Chinese writing with wiki among Chinese upper primary school students. *Computers and Education*, *126*, 359–375. https://doi.org/10.1016/j.compedu.2018.06.009

Liu, Gi-Zen, Rahimi, Masoud, & Fathi, Jalil. (2022). Flipping writing metacognitive strategies and writing skills in an English as a foreign language collaborative writing context: A mixed-methods study. *Journal of Computer Assisted Learning*, *38*(6), 1730-1751.

Liu, S. H.-J., Lan, Y.-J., & Ho, C. Y.-Y. (2014). Exploring the Relationship between Self Regulated Vocabulary Learning and Web-Based Collaboration. *Educational Technology & Society*, *17*(4), 404–419.

Long, Michael H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*(2), 126-141.

Long, Michael H. (1985). Input and second language acquisition theory. In S.M. Gass and C.G. Madden (eds) *Input in second language acquisition* (pp. 377–393). Rowley, NA: Newbury House.

Long, Michael H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R.B. Ginsberg and C. Kramsch (eds) *Foreign language research in a cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

Long, Michael H. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie and T.K. Bhatia (eds) *Handbook of language acquisition. Vol. 2: Second language acquisition* (pp. 413–468). New York: Academic Press.

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

Lund, A. (2008). Wikis: a collective approach to language production. *ReCALL (Cambridge, England)*, *20*(1), 35–54. https://doi.org/10.1017/S0958344008000414

Ma, Q. (2020). Examining the role of inter-group peer online feedback on wiki writing in an EAP context. *Computer Assisted Language Learning*, *33*(3), 197–216. https://doi.org/10.1080/09588221.2018.1556703

Ma, Qing. (2019). Examining the role of inter-group peer online feedback on wiki writing in an EAP context. *Computer Assisted Language Learning*, *33*(3), 197-216.

Mak, B., & Coniam, D. (2008). Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System (Linköping)*, *36*(3), 437–455. https://doi.org/10.1016/j.system.2008.02.004

Malmqvist, Anita. (2005). How Does Group Discussion in Reconstruction Tasks Affect Written Language Output? *Language Awareness*, *14*(2-3), 128-141.

Manchón, R. (2011b) Writing to learn the language: Issues in theory and research. In R. Manchón (ed.) *Learning-to-write and writing-to-learn in an additional language* (pp. 61–84). Amsterdam/Philadelphia: John Benjamins.

Mohammadi, Kourosh, Jafarpour, Aliakbar, Alipour, Javad, & Hashemian, Mahmood. (2023). The Impact of Different Kinds of Collaborative Prewriting on EFL Learners' Degree of Engagement in Writing and Writing Self-Efficacy. *Reading & Writing Quarterly*, *Ahead-of-print* (Ahead-of-print), 1-18.

Nassaji, Hossein, & Jun Tian. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research: LTR*, *14*(4), 397-419.

Neumann, Heike, & McDonough, Kim. (2015). Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing. *Journal of Second Language Writing*, *27*(Mar), 84-104.

Niu, Ruiying, Jiang, Lin, & Deng, Yuan. (2018). Effect of Proficiency Pairing on L2 Learners' Language Learning and Scaffolding in Collaborative Writing. *The Asia-Pacific Education Researcher*, *27*(3), 187-195.

Nordlof, J. (2014). Vygotsky, Scaffolding, and the Role of Theory in Writing Center Work. *The Writing Center Journal*, *34*(1), 45–64. https://doi.org/10.7771/2832-9414.1785

Ortega, L. (2009). Interaction and attention to form in L2 text-based computer-mediated communication. In A. Mackey and C. Polio (eds) *Multiple perspectives on interaction. Second language research in honour of Susan Gass* (pp. 226–253). New York: Routledge.

Ortoleva, Giulia, & Bétrancourt, Mireille. (2015). Collaborative writing and discussion in vocational education: Effects on learning and self-efficacy beliefs. *Journal of Writing Research*, *7*(1), 95-122.

Oxford, Rebecca L. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal (Boulder, Colo.)*, *81*(4), 443-456.

Pajares, F., & Valiante, G. (2001). Response format in self-efficacy: Greater discrimination increases prediction. *Counseling and Development*, *33*(4), 35-43.

Pajares, Frank. (2003). SELF-EFFICACY BELIEFS, MOTIVATION, AND ACHIEVEMENT IN WRITING: A REVIEW OF THE LITERATURE. *Reading & Writing Quarterly*, *19*(2), 139-158.

Pajares, Frank. (2007). Empirical Properties of a Scale to Assess Writing Self-Efficacy in School Contexts. *Measurement and Evaluation in Counseling and Development*, *39*(4), 239-249.

Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, *31*(1), 117–134. https://doi.org/10.1177/0267658314536435

Pallotti, G. (2021). Measuring Complexity, Accuracy, and Fluency (CAF). In *The Routledge Handbook of Second Language Acquisition and Language Testing* (1st ed., Vol. 1, pp. 201–210). Routledge. https://doi.org/10.4324/9781351034784-23

Platt, E., & Brooks, F. B. (2002). Task Engagement: A Turning Point in Foreign Language Development. *Language Learning*, *52*(2), 365–400. https://doi.org/10.1111/0023 8333.00187

Qiu, Xuyan, & Lee, Man-Kit. (2020). Regulated learning and self-efficacy beliefs in peer collaborative writing: An exploratory study of L2 learners' written products, task discussions, and self-reports. *System (Linköping)*, *93*, 102312.

Rahimi, Masoud, & Fathi, Jalil. (2022). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self efficacy: A mixed methods study. *Computer Assisted Language Learning*, *35*(9), 2627 2674.

Reinders, Hayo. (2009). Learner uptake and acquisition in three grammar-oriented production activities. *Language Teaching Research: LTR*, *13*(2), 201-222.

Reinhardt, Jonathon. (2019). Social media in second and foreign language teaching and learning: Blogs, wikis, and social networking. *Language Teaching*, *52*(1), 1-39.

Roed, Jannie. (2003). Language Learner Behaviour in a Virtual Environment. *Computer Assisted Language Learning*, *16*(2-3), 155-172.

Schnaubert, Lenka, & Bodemer, Daniel. (2022). Group awareness and regulation in computer supported collaborative learning. *International Journal of Computer-supported Collaborative Learning*, *17*(1), 11-38.

Schrage, M. (1994). Writing to collaborate; collaborating to write. In J.S. Leonard, C.E. Wharton, R.M. Davis and J. Harris (eds) Authority and textuality. *Current views of collaborative writing* (pp. 17–24). West Cornwall, CT: Locus Hill Press.

Schunk, Dale H., & Zimmerman, Barry J. (2007). Influencing Children's Self-Efficacy and Self-Regulation of Reading and Writing Through Modeling. *Reading & Writing Quarterly*, *23*(1), 7-25.

Selcuk, Hasan, Jones, Jane, & Vonkova, Hana. (2021). The emergence and influence of group leaders in web-based collaborative writing: Self-reported accounts of EFL learners. *Computer Assisted Language Learning*, *34*(8), 1040-1060.

Sever, Esra, & Akyol, Hayati. (2022). The Impact of Collaborative Learning Techniques on Written Expression, Self- Regulation and Writing Motivation. *International Electronic Journal of Elementary Education*, *14*(5), 587-603.

Sexton, Melissa, Harris, Karen R., & Graham, Steve. (1998). Self-Regulated Strategy Development and the Writing Process: Effects on Essay Writing and Attributions. *Exceptional Children*, *64*(3), 295-311.

Shehadeh, Ali. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, *20*(4), 286-305.

Slavin, Robert E. (1991). *Student Team Learning: A Practical Guide to Cooperative Learning*. Third Edition.

Stahl, Gerry. (2006). *Group Cognition - Computer Support for Building Collaborative Knowledge*. Michigan: The MIT Press.

Storch, N. (2001). Comparing ESL learners' attention to grammar on three different classroom tasks. *RELC Journal*, *32*(2), 104-124.

Storch, Neomy, & Aldosari, Ali. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research:LTR*, *14*(4), 355-375.

Storch, Neomy, & Aldosari, Ali. (2013). Pairing learners in pair work activity. *Language Teaching Research: LTR*, *17*(1), 31-48.

Storch, Neomy. (1998). Comparing Second Language Learners' Attention to Form Across Tasks. *Language Awareness*, *7*(4), 176-191.

Storch, Neomy. (2004). Using activity theory to explain differences in patterns of dyadic interactions in an ESL class. *Canadian Modern Language Review*, *60*(4), 457-480.

Storch, Neomy. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, *14*(3), 153-173.

Storch, Neomy. (2013). *Collaborative Writing in L2 Classrooms (New Perspectives on Language and Education)*. Blue Ridge Summit, PA.

Storch, Neomy. (2018). Collaborative writing. *The TESOL Encyclopedia of English Language Teaching*, 1-6.

Su, Fan, & Zou, Di. (2022). Technology-enhanced collaborative language learning: Theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, *35*(8), 1754-1788.

Su, You, Li, Yanyan, Liang, Jyh-Chong, & Tsai, Chin-Chung. (2019). Moving literature circles into wiki-based environment: The role of online self-regulation in EFL learners' attitude toward collaborative learning. *Computer Assisted Language Learning*, *32*(5-6), 556-586.

Su, You, Li, Yanyan, Liang, Jyh-Chong, & Tsai, Chin-Chung. (2019). Moving literature circles into wiki-based environment: The role of online self-regulation in EFL learners' attitude toward collaborative learning. *Computer Assisted Language Learning*, *32*(5-6), 556 586.

Su, You, Liang, Jyh-Chong, Zheng, Chunping, & Tsai, Chin-Chung. (2023). Co-regulation strategies and their associations with writing self-efficacy in a computer-mediated collaborative writing setting. *Journal of Second Language Writing*, *59*, 100972.

Su, You, Zheng, Chunping, Liang, Jyh-Chong, & Tsai, Chin-Chung. (2018). Examining the relationship between English language learners' online self-regulation and their self efficacy. *Australasian Journal of Educational Technology*, *34*(3).

Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H.G. Widdowson* (pp. 125–144).Oxford, UK: Oxford University Press.

Swain, Merrill, & Lapkin, Sharon. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, *37*(3), 285 304.

Swain, Merrill, & Lapkin, Sharon. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, *37*(3), 285 304.

Swain, Merrill. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. M. Gass and C. G. Madden (eds) *Input in second language acquisition* (pp. 235–256). Rowley, MA: Newbury House.

Swain, Merrill. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, *50*(1), 158–164.

Swain, Merrill. (1995). Three functions of output in second language learning. In G. Crook and B. Seidlhofer (eds) *Principles and practices in applied linguistics: Studies in honour of H. Widdowson* (pp. 125–144). Oxford: Oxford University Press.

Swain, Merrill. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (ed.) *Sociocultural theory and second language learning* (pp. 97–114). Oxford: Oxford University Press.

Swain, Merrill. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (ed.) *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95–108). London, UK: Continuum.

Swain, Merrill. (2010). Talking-it-through: Languaging as a source of learning. In R. Batestone (ed.) *Sociocognitive perspectives on language use and language learning* (pp. 112 130). Oxford: Oxford University Press.

Tai, Hung-Cheng. (2016). Effects of Collaborative Online Learning on EFL Learners' Writing Performance and Self-Efficacy. *English Language Teaching (Toronto)*, *9*(5), 119.

Teng, Feng. (2020). Young learners' reading and writing performance: Exploring collaborative modeling of text structure as an additional component of self-regulated strategy development. *Studies in Educational Evaluation*, *65*, 100870.

Teng, L., & Zhang, L. J. (2018). Effects of motivational regulation strategies on writing performance: A mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacognition and Learning*, *13*(2), 213-240.

The Curriculum Development Council. (2017). *English Language Education Key Learning Area Curriculum Guide (Primary 1 - Secondary 6)*. Education Bureau.

Tocalli-Beller, Agustina. (2003). Cognitive conflict, disagreement and repetition in collaborative groups: Affective and social dimensions from an insider's perspective. *Canadian Modern Language Review*, *60*(2), 143.

Tsai, Pi-hua. (2019). Beyond self-directed computer-assisted pronunciation learning: A qualitative investigation of a collaborative approach. *Computer Assisted Language Learning*, *32*(7), 713-744.

Villarreal, Izaskun, & Gil-Sarratea, Nora. (2020). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research: LTR*, *24*(6), 874-897.

Villarreal, Izaskun, & Gil-Sarratea, Nora. (2020). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research: LTR*, *24*(6), 874-897.

Vygotskiĭ, L., & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Cambridge.

Vygotskiĭ, L., & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Cambridge.

Wajnryb, R. (1990). *Grammar dictation* (Resource books for teachers). Oxford: Oxford University Press.

Wang, Y. huei, & Liao, H. C. (2017). Learning performance enhancement using computer assisted language learning by collaborative learning groups. *Symmetry (Basel)*, *9*(8), 141-. https://doi.org/10.3390/sym9080141

Wang, Y.-C. (2014). Using wikis to facilitate interaction and collaboration among EFL learners: A social constructivist approach to language teaching. *System (Linköping)*, *42*(1), 383 390. https://doi.org/10.1016/j.system.2014.01.007

Wang, Yu-Chun. (2015). Promoting collaborative writing through wikis: A new approach for advancing innovative and active learning in an ESP context. *Computer Assisted Language Learning*, *28*(6), 499-512.

Warschauer, Mark, & Grimes, Douglas. (2007). AUDIENCE, AUTHORSHIP, AND ARTIFACT: THE EMERGENT SEMIOTICS OF WEB 2.0. *Annual Review of Applied Linguistics*, *27*, 1-23.

Watanabe, Yuko. (2008). Peer-Peer Interaction between L2 Learners of Different Proficiency Levels: Their Interactions and Reflections. *Canadian Modern Language Review*, *64*(4), 605-635.

Wells, Gordon. (1999). Using L1 to master L2: A response to Antón and DiCamilla's "socio cognitive functions of L1 collaborative interaction in the L2 classroom". *The Modern Language Journal (Boulder, Colo.)*, *83*(2), 248-254.

Wigglesworth, Gillian, & Storch, Neomy. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, *26*(3), 445-466.

Wigglesworth, Gillian, & Storch, Neomy. (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing*, *21*(4), 364-374.

Wood, David, Bruner, Jerome S., & Ross, Gail. (1976). THE ROLE OF TUTORING IN PROBLEM SOLVING. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100.

Xie, Kui, Hensley, Lauren C., Law, Victor, & Sun, Zhiru. (2019). Self-regulation as a function of perceived leadership and cohesion in small group online collaborative learning. *British Journal of Educational Technology*, *50*(1), 456-468.

Xu, Y., Park, H., & Baek, Y. (2011). A New Approach Toward Digital Storytelling: An Activity Focused on Writing Self-efficacy in a Virtual Learning Environment. *Educational Technology & Society*, *14*(4), 181–191.

Yang, L. (2014). Examining the mediational means in collaborative writing: Case studies of undergraduate ESL students in business courses. *Journal of Second Language Writing*, *23*, 74-89.

Yang, Luxin. (2014). Examining the mediational means in collaborative writing: Case studies of undergraduate ESL students in business courses. *Journal of Second Language Writing*, *23*(Mar), 74-89.

Yeh, Shiou-Wen, & Chen, Cheng-Ting. (2019). EFL Learners' Peer Negotiations and Attitudes in Mobile-Assisted Collaborative Writing. *Language Education & Assessment*, *2*(1), 41.

Zhang, Meixiu, & Chen, Wenting. (2022). Assessing collaborative writing in the digital age: An exploratory study. *Journal of Second Language Writing*, *57*, 100868.

Zhang, Meixiu, Gibbons, John, & Li, Mimi. (2021). Computer-mediated collaborative writing in L2 classrooms: A systematic review. *Journal of Second Language Writing*, *54*, 100854.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82-91.

Zimmerman, B. J., & Martinez-Pons, M. (1992). Perceptions of efficacy and strategy use in the self-regulation of learning. In D. H. Schunk & J. Meece (Eds.), *Student perceptions in the classroom: Causes and consequences* (pp. 185–207).Hillsdale, NJ: Erlbaum.

Zimmerman, Barry J., & Bandura, Albert. (1994). Impact of Self-Regulatory Influences on Writing Course Attainment. *American Educational Research Journal*, *31*(4), 845-862.

Zimmerman, Barry J., & Risemberg, Rafael. (1997). Becoming a Self-Regulated Writer: A Social Cognitive Perspective. *Contemporary Educational Psychology*, *22*(1), 73-101.

Appendix A: Summary of Previous Studies

Studies	Context	Tasks	Implementation	Outcomes & Implications	limitations
Fortune (2005)	N= 8 L2= English Advanced University students in the UK	Dictogloss	9 weeks Pairs Data collection / analysis: 1. Follow-up interviews 2. Transcripts 3. Language-related episodes analysis	1. The advanced learners used metalanguage much more often than their intermediate counterparts, particularly in lexical episodes in dictogloss tasks	1. The sample size was small 2. A short-term study cannot capture changes / developments over time
Abadikhah (2012)	N= 36 English as a foreign language (EFL) Low intermediate University students in Iran	Mechanical exercises (substitution) Description Dictogloss	8 weeks Completed mechanical exercises in 15 mins 2 groups Data collection / analysis: 1. Pre-test and post-test writings	1. The results revealed significant improvement for both groups from the pre-test to the post-test of mechanical and meaningful tasks ($p < .001$) 2. There was a significant relationship between	1. The sample size was small à it was difficult to examine possible interaction effects of high and low scoring and practice type

			2. The collaborative dialogue of the participants was tape-recorded and transcribed 3. Language-related episodes analysis	task type and number of language-related episodes (LREs) 3. Meaning-focused tasks yielded more LREs than form-focused tasks	2. The pre-test and post-test tasks only targeted one grammatical form (relative clauses)
Dobao (2012)	N= 111 L1: English Foreign language= Spanish (intermediate) University students in the USA 6 intact classes	Jigsaw task (story based on a set of pictures)	Seven weeks Completed a jigsaw task in 30 mins Pairs (n=15), small groups (n=15) of 4 students, 21 students working individually Data collection / analysis: 1. 15 texts (groups), 15 texts (pairs) and 21 texts (individuals) 2. 722 mins audio-recorded oral interaction	1. No statistically significant difference in terms of complexity among individuals, pairs and groups 2. CW can enhance the linguistic accuracy of written texts 3. Groups > pairs > individuals, in terms of producing grammatically accurate texts	1. Only focused on intermediate level learners 2. Only one type of CW task was employed à cannot be generalized across task types or proficiency levels

			3. Measured learners' writing performance in terms of CFA 4. Coding language-related episodes		
Wang (2015)	N= 48 L1= Mandarin L2= English University students in Taiwan	A business letter	One semester: business writing lessons Wrote a business letter in 60 mins The tests were verified for content validity by two experts Data collection / analysis: 1. Writing tests 2. A survey	1. Statistically significant differences were found in wiki-based writing in terms of promoting students' interest in language learning and boosting students' writing competencies	1. A small number of university students in Southern Taiwan àthe findings may not be generalizable to other educational settings / backgrounds 2. Participants' English level was not specified
Law, Ge & Eseryel (2016)	N= 131 L1= English University students in the USA	Jigsaw task (discuss an ill-structured problem)	Completed a collaborative problem-solving task (for 20 - 30 mins) Data collection / analysis: 1. Self-regulation and problem-solving scales	1. A four-factor model of self-regulation clarification and resolution, elaboration, refuting, and summarization occurred in CW	1. The study was conducted in a lab environment (instead of a naturalistic environment) where students

				2. Self-regulation in CW is not a uni-directional process, but a bi-directional process	were assigned to ad hoc teams 2. Self-report instrument may not be able to represent participants' real intention and experiences
Qiu & Lee (2020)	N= 24 L1= Mandarin L2= English (intermediate) EFL university students in China	Rewriting an ending	Rewrote an ending of at least 200 words in one hour Data collection / analysis: 1. Pre-test, post-test and delayed-post-test writings 2. Task discussions 3. Reflective notes 4. Follow-up interviews	1. CW helped students improve their writing in terms of CFA 2. Students had higher self-efficacy and self-regulation for L2 writing	1.The sample size was small 2. The participants' English proficiency and levels of writing self-efficacy were not measured before the CW implementation

Rahimi & Fathi (2021)	N= 67 EFL students in Iran	Two writing tasks	One semester: 12 ninety-min sessions Oxford Placement Test was adopted to measure students' English proficiency Experimental group (N=35) Control group (N= 32) Two 50-min writing tasks Data collection / analysis: 1. Microgenetic method 2. An individual semi-structured interview 3. SLWS and SLWSS 4. Language-related episodes	1. Significant differences between the wiki-mediated and non-wiki collaborative writing groups in improving the students' writing performance, writing self-regulation and writing self-efficacy 2. Wiki-mediated CW could enhance students' writing performance in terms of content, organization, and language use, and their SRSE	1. The study did not interview the participants in the non-wiki collaborative group to check their perceptions and attitudes 2. The interview was taken place on the last day of the course, which could have influenced participants' responses to the interview questions 3. Longitudinal research might be needed to explore students' improvement in writing performance, self-regulation and self-efficacy as this current

					study only lasted one term

Appendix B: Second language writing self-regulation scale

Items	Never	Rarely	Sometimes	Often	Always
1. I understand how to alleviate the stress from learning to write in English.					
2. I have special techniques to accomplish my learning objectives for writing in English.					
3. I feel satisfied with my own special methods for reducing the stress of writing in English.					
4. I have particular techniques to maintain my concentration when learning to write in English.					
5. I persevere until I achieve the goals I set for myself when learning to write in English.					
6. I believe I can achieve my goals faster than expected when learning to write in English.					
7. I can cope with the stress from learning writing in English immediately.					

8. When it comes to learning writing in English, I think my methods of controlling procrastination are effective.					
9. I know how to arrange the environment to make learning more efficient when learning writing in English.					

Adapted from Han and Hiver (2018)

Appendix C: Second language writing self-efficacy scale

Items	Never	Rarely	Sometimes	Often	Always
1. I have confidence in my ability to write in English.					
2. I possess the knowledge to write well in English.					
3. My writing in English adheres to a logical organizational structure.					
4. With sufficient effort, I am certain I can become adept at writing in English.					
5. I can write essays that are relevant and appropriate to the assignment.					
6. When writing in English, I can accurately and effectively present my viewpoint or arguments.					
7. I am assured I can succeed in writing lessons even if they are difficult.					

Adapted from Han and Hiver (2018)

Appendix D: Interview questions

1. Did you enjoy the collaborative writing course on Google Docs.? If yes, would you please talk about your feelings about the lessons?

2. Can you identify specific areas of your writing that have improved as a result of technology-mediated collaborative writing? For instance, grammar, vocabulary, organization, or others?

3. Did the collaborative writing lessons help you experience less negative feelings, such as anxiety, lack of self-confidence, fear of evaluation, etc., when you wrote in English?

4. To what extent do you feel that technology-mediated collaborative writing has affected your belief in your ability to tackle L2 writing tasks?

5. How does the feedback from your peers during the collaborative writing process impact your self-assurance as an L2 writer?

6. How do you plan, monitor, and evaluate your writing when working collaboratively?

7. In what ways has working with others helped you to become more disciplined or self-directed in your writing practices?

8. Do you think that the collaborative writing lessons should be included in normal English writing classrooms? Please elaborate on your answer.

9. How different was the collaborative writing lessons from normal English writing lessons courses you had experienced in other classes?