Journal of Second Language Writing 54 (2021) 100854

Contents lists available at ScienceDirect

Journal of Second Language Writing

journal homepage: www.elsevier.com/locate/jslw

Computer-mediated collaborative writing in L2 classrooms: A systematic review

Meixiu Zhang [a,*], John Gibbons [b], Mimi Li [b]

[a] Department of Classical and Modern Languages and Literatures, Texas Tech University, United States
[b] Department of Literature and Languages, Texas A&M University-Commerce, United States

ARTICLE INFO

Keywords:
Computer-mediated communication
Collaborative writing
Second language writing
Methodological synthesis
Systematic review
Scoping review

ABSTRACT

Fueled by a keen awareness of the affordances of technology and the predominance of the communicative approach, computer-mediated collaborative writing (CMCW) has gained fast-growing interest in second language (L2) contexts. To illuminate research foci and guide future research efforts in this burgeoning area, the current study provided a substantive and methodological review of 113 primary studies on CMCW in L2 contexts, covering qualitative, quantitative, and mixed-methods research. Each study was coded for substantive features, including research context, study setup, components analyzed, and the analytic frameworks. We also coded the sample for methodological practices that were applicable to all research (e.g., instrument development, reliability estimates) and those related to a specific research orientation (e.g., use of statistical analyses, validity strategies). The results indicated (a) a strong preference for meaning-focused writing tasks covering various genres, (b) considerable heterogeneity in the metrics for measuring texts, (c) limited interests in L2 learners' attention to form. While there was a strength in reporting certain strategies for enhancing research validity, the results revealed methodological concerns such as insufficient reporting of learner/context information, reliability estimates, effect sizes, and exact *p* values. Based upon the findings, we provided a number of empirically grounded suggestions for future research.

1. Introduction

Driven by substantial research evidence supporting the value of learner interaction and collaboration in L2 development (Kim, 2008; Shehadeh, 2011), collaborative tasks have been widely implemented in language classrooms. Collaborative writing (CW), defined as a pedagogical task in which two or more learners interact with each other throughout the writing process to co-produce one text, has gained L2 researchers' growing attention in the past decades (Storch, 2019). In a recent review, Lei and Liu (2019) reported that applied linguists' interest in CW increased by over seven times from 2005 to 2016. Prior research indicates that CW benefits L2 learning as it promotes the co-construction of L2 lexical and grammatical knowledge (Fernández Dobao, 2012; Zhang & Crawford, 2021), uptake of corrective feedback (Wigglesworth & Storch, 2012), the enhancement of audience awareness (Storch, 2013), and L2 writing development (Shehadeh, 2011).

One important driving force behind the thriving of CW research relates to the incorporation of computer-mediated communication (CMC) in CW. The properties of CMC, which includes "synchronous and asynchronous communication capacity, multi-way

* Corresponding author.
E-mail addresses: Meixiu.Zhang@ttu.edu (M. Zhang), John.Gibbons@tamuc.edu (J. Gibbons), mimi.li@tamuc.edu (M. Li).

https://doi.org/10.1016/j.jslw.2021.100854
Received 9 July 2021; Received in revised form 18 October 2021; Accepted 25 October 2021
Available online 2 December 2021

communication, and high interactivity" (Luppicini, 2007, p.142), greatly contribute to the utility of CMCW in language classrooms. Web 2.0 technologies such as Google Docs and wikis afford ample opportunities for learners to communicate, jointly write, and mutually scaffold L2 learning beyond time constraints of onsite classrooms (Li, 2018). A growing body of empirical research has explored the potential of CMCW in diverse settings over the past decades (e.g., Mak & Coniam, 2008; Rogers, 1998; Torres & Cung, 2019). This burgeoning area has employed an array of research approaches (e.g., case study, grounded theory, quasi-experimental), examined diverse digital tools (e.g., wikis, Google Docs, blogs), developed principled frameworks for analyzing learner data, and investigated a wide range of factors, such as task modality, task type, learner proficiency, and learner agency (Abrams, 2019; Arnold, Ducate, & Kost, 2012; Li & Kim, 2016; Li & Zhu, 2017a). As research interests continue to grow, there is a clear need for L2 researchers to better understand research foci, analytic frameworks, and research practices in this area of inquiry. Thus, this study, as a sequel to Zhang and Plonsky (2020) who systematically reviewed face-to-face CW research, aims to provide a substantive and methodological review of empirical L2 research on CMCW. Of particular note, this study was a scoping review that provides "an overview of a broad field" instead of a synthesis of research findings (Moher, Stewart, & Shekelle, 2015, p.1). In the following section, we first review the theoretical and empirical support for CMCW, and then describe the L2 features and aspects that are often analyzed in this area. Last, we discuss methodological considerations in this domain and present the research questions (RQs).

2. Background

2.1. Theoretical support for computer-mediated collaborative writing

With the aid of technological tools, CMCW tasks weld together two elements that are critical for language learning in technology-supported environments: peer interaction and written production in the L2 (Abrams, 2016). The role of peer interaction and the act of producing meaningful texts in the L2 in CMC settings, though conceptualized differently, have been underscored in both sociocultural theory (Vygotsky, 1978) and the interactionist approach to L2 learning (Long, 1983) The sociocultural theory posits that language development is a result of social interaction during which language uses serve dual functions: a means for communication and a cognitive tool for mediating the co-construction of L2 knowledge (Antón & DiCamilla, 1998; Swain & Lapkin, 1998; Vygotsky, 1978). Computer-supported collaborative dialogue during CW provides a site for learners to elicit and provide scaffolding, pool resources together to resolve linguistic issues, and collectively construct L2 knowledge (Donato, 1994; Li & Kim, 2016). From an interactionist perspective, computer-mediated peer interaction affords modified input, corrective peer feedback, and opportunities for producing modified output, which promotes L2 development (Smith, 2004; Ziegler, 2016). Additionally, the act of co-producing an L2 text in CMC settings, whether conceptualized as a means for mediating the L2 learning processes or a process that generates meaningful written output, is deemed highly valuable for L2 development (Abrams, 2016; Storch, 2019; Williams, 2012).

2.2. Empirical research on computer-mediated collaborative writing

Informed by the above-mentioned theoretical frameworks, researchers have empirically surveyed the potential of CMCW in language learning. One strand of research primarily employed quasi-experimental studies to explore the effects of CMCW on writing gains and other aspects of learning. For instance, Bikowski and Vithanage (2016) reported significant positive effects of Google Docs-based CW on ESL learners' development of L2 writing proficiency over one semester. Additional encouraging evidence was found in Hsu and Lo (2018) who reported that English as a Foreign Language (EFL)dyads of learners improved significantly more in both the content quality and linguistic accuracy of their writing than individual writers. Similarly, Wang (2015) examined English for specific purposes (ESP) learners and found that CMCW significantly enhanced learners' business writing skills, interests in L2 learning, and collaboration skills.

Another series of experimental studies have investigated the effects of CMCW on the written products and interaction in comparison with individual writing or face-to-face (F2F) CW (see Elabdali, 2021 for a meta-analysis). For instance, focusing on German learners, Strobl (2014) found no difference in the complexity, accuracy, and fluency of the texts between the individual writing group and the CMCW group. However, she captured positive effects of CMCW on content selection and organization of the written products. Comparing CMCW and F2F CW, Wu (2015) indicated no advantages of blog-based CW over traditional CW in enhancing the EFL learners' written products.

Aside from experimental studies, qualitative and quantitative descriptive research constitutes a large portion of existing literature, which primarily falls under two strands. The first strand aimed to depict the different components (e.g., interaction, revision, written products) and behaviors (e.g., strategy use) involved in CMCW tasks. For instance, multiple studies have employed fine-grained analyses to understand collaboration patterns formed in Google-Docs-based CW (Abrams, 2016) and wiki-based CW (Ducate, Anderson, & Moreno, 2011; Li & Kim, 2016; Li & Zhu, 2013). In terms of learner behaviors, researchers have examined writing strategies (Kost, 2011), politeness strategies (Li, 2012; Li & Zhu, 2013), and behaviors indicative of collaborative autonomous learning (Kessler & Bikowski, 2010).

The other strand accounted for the link among the different components of CMCW in relation to learner factors or task factors. For instance, Abrams (2019) and Li and Zhu (2017b) explored the link between collaboration patterns and learners' co-constructed texts. Cho (2017) probed how the dynamics of collaboration was influenced by the mode of interaction and learner perception in synchronous CMCW, whereas Selcuk, Jones, and Vonkova (2019) looked into the influences of group leaders on Turkish EFL students' learning and task completion process in Facebook-based CW.

To summarize, prior empirical research has provided convincing evidence for the facilitative role of CM- CW in L2 learning,

although its benefits on the written products remain inconclusive. Also noteworthy is the diversity in multiple aspects of the research area, which was mirrored in the various research methodologies employed, the different research questions pursued, the wide range of technical tools and platforms surveyed, as well as the various analyses undertaken. Such a diversity points to a pressing need for a systematic review of prior research foci, analytic frameworks, and research practices. The following section reviews what and how elements and processes of CMCW were typically analyzed in prior investigations.

2.3. Typical analyses in computer-mediated CW

To understand the process, product, and learners' perspective, previous research has mainly analyzed five relevant aspects of CMCW, including learner interaction, revisions, the written products, dynamics of collaboration, and task perception. To begin with, learner interaction in CMCW may take place in the form of online chats, discussion posts, and/or comment records. It has been analyzed for three major purposes: (a) detecting the level of engagement (e.g., Alghasab & Handley, 2017; Li & Zhu, 2017a, 2017b), (b) describing the focus of the interaction to understand what aspects of the task have been prioritized (e.g., Elola & Oskoz, 2010), and (c) providing insight on learners' attention to form (e.g., Díez-Bedmar & Pérez-Paredes, 2012) through analyzing language-related episodes (LREs). LREs, defined as any part of interaction where learners discuss 'the language they are producing, question their language use, or correct themselves or others' (Swain & Lapkin, 1998, p.326), are believed to represent opportunities for L2 learning (Swain & Lapkin, 1998, 2013).

Concerning the analysis of revisions, the main approaches were the framework of writing change functions (Li & Zhu, 2013; Li & Kim, 2016; Mak & Coniam, 2008) and the taxonomy of revisions (Faigley & Witte, 1981). The former distinguished revision acts such as adding, deleting, rephrasing, reorganizing, as well as revisions of self-generated texts and those of other-produced texts. The latter considered both acts of revision (e.g., addition, deletion, substitution) and the linguistic foci of the revisions (e.g., lexical revisions, word order, tense) (Arnold et al., 2012; Lai, Lei, & Liu, 2016).

Resembling face-to-face CW and other L2 writing research (Johnson, 2017; Zhang & Plonsky, 2020), written products in CMCW were often examined via text quality and the complexity, accuracy, and fluency (CAF) framework (Housen & Kuiken, 2009). Researchers have also examined the dynamics of collaboration to afford additional insight on learners' collaboration process in technology-supported CW, mostly drawing upon a model developed by Storch (2002),201) which accounts for the extent to which learners contribute equally and engage mutually. In addition, efforts have been undertaken to gauge learners' perception of CMCW tasks through methods such as interviews and questionnaires (e.g., Li & Zhu, 2013; Strobl, 2014).

To provide a clear and comprehensive picture of the key concepts and analyses underpinning CMCW research, this study aims to systematically review how the above-mentioned components were examined in prior empirical studies.

2.4. Methodological considerations in empirical L2 research

Aside from the substantive review, this paper also attempts to review research and reporting practices in this domain, in response to the recent calls for promoting methodological rigor in empirical L2 research and transparency in reporting research findings (e.g., Al-Hoorie & Vitta, 2019; Plonsky & Derrick, 2016). As noted in Byrnes (2013), "methodologies no longer have ancillary status in our work" (p.825). Recently, in a review of F2F CW research, Zhang and Plonsky (2020) identified multiple issues which might undermine research validity and obscure the interpretation of research findings in this area: (a) insufficient reporting of descriptive statistics and reliability estimates of data coding, (b) limited piloting of instruments, and (c) inconsistent implementation of pre-task training. Nevertheless, it remains unknown whether such issues were shared by CMCW research. Therefore, an additional goal of this study was to review research and reporting practices in CMCW research with the hope to better direct future research efforts in this promising area. Of particular note, while Zhang and Plonsky (2020) focused on quantitative research, the current study included qualitative research because this research orientation is integral to the CMCW research. Different methodological aspects were reviewed for research of different orientations (See Section 3.2 for details). Specifically, this study aims to answer the following RQs:

1 What are the study contexts and study setup features in empirical L2 CMCW research?
2 What are the methodological design features in empirical L2 CMCW research?
3 What are the theoretical frameworks of CMCW research and the different components that have been analyzed? How have they been analyzed?
4 What research and reporting practices have been adopted in empirical L2 CMCW research?

3. Methods

3.1. Identification and retrieval of primary studies

While many recent substantive and methodological reviews focused exclusively on quantitative L2 research (e.g., Plonsky, 2013, 2014; Zhang & Plonsky, 2020), we included both qualitative and quantitative studies in order to provide a comprehensive picture of this domain. Also, both published and unpublished studies were included to help alleviate publication bias (Rosenthal, 1979). Specifically, we included CW studies that met the following criteria: (a) empirical research, (b) conducted in L2 contexts, (c) implemented in CMC environments, and (d) published in English before June 2020.

To retrieve relevant studies, we used a combination of the keywords "collaborative writing", AND "language", AND "computer" OR

"technology" OR "online" to search six databases, including Education Resources Information Center, Linguistic and Language Behavior Abstracts, ProQuest, PsycArticles, PsycInfo, and Linguistics Abstracts Online. We also searched the first 300 entries in Google Scholar. An examination of the abstract and the method section led to 115 studies that met the inclusion criteria. Two dissertations were inaccessible and therefore excluded, which resulted in a sample of 113 studies, including 83 journal articles, eleven Ph.D. dissertations, eight conference proceedings, six M.A. theses, and five book chapters. Fig. 1 visually presented the distribution of the 113 studies over time, revealing a robust growth of research interests in the past decade (Tables 1 and 2).

3.2. Coding

Following Zhang and Plonsky (2020), a coding scheme (see Table 3) was developed to extract five types of information: (a) study context, (b) study design, (c) study setup, (d) analyses of L2 features, and (e) research and reporting practices.

Based on our data set, we distinguished quasi-experimental (i.e., studies involving a treatment), survey (i.e., studies aiming to describe traits of a population by surveying a sample of the group), and descriptive (i.e., studies relying on numeric data without treatments) under quantitative research approaches, informed by Dörnyei (2007) and Brown (2005). We also defined the following qualitative approaches: (a) case study (i.e., research examining a small number of participants' behaviors and perspectives intensively over a period of time); (b) action research (i.e., research addressing practical issues in one's own context of work), and (c) grounded theory (i.e., research that derives theoretical constructs inductively from empirical data), informed by Heigham and Croker (2009). Given that CMC settings afford multiple task modalities during the collaboration process, we included mode of collaboration as a study setup feature in the coding scheme, distinguishing synchronous (i.e., real-time collaboration), asynchronous (i.e., delayed-time collaboration), and hybrid collaboration (i.e., including both real-time and delayed-time collaboration).

Regarding research and reporting practices, we retrieved information about practices that were relevant to all studies in our sample, including instrument development, pre-task training, and reporting of reliability coefficients. To capture the distinctive features of quantitative and qualitative research, we also designated separate sets of items. For quantitative studies, we retrieved information about statistical analyses and features associated with researcher transparency in quantitative L2 research (e.g., reporting of effect sizes, *p* values, and assumption checking) (see Larson-Hall & Plonsky, 2015 for further discussion). For qualitative studies, we examined four validity strategies that were believed to be relevant to most qualitative research (Creswell & Creswell, 2017): (a) triangulation: utilizing multiple sources of data or groups of participants to justify themes (Richards, 2009), (b) member checking: presenting 'polished' or 'semi-polished' products to participants and provide them an opportunity to comment on the findings (Creswell & Creswell, 2017), (c) a thick description: providing a description that "gives the context of an act, states the intentions and meanings that organize the action, traces the evolution and development of the act, and presents the action as a text that can then be interpreted" (Denzin, 1989, p. 30, c.f. Ponterotto, 2006), and (d) researcher bias clarification: a reflective account of potential researcher bias, position, or subjectivity that may influence the interpretation of findings (Richards, 2009).

The coding scheme was piloted by the first and third authors, which resulted in revisions in subcategories and operational definitions for multiple items. Once the coding scheme was refined, the second author undertook extensive training and worked with the first author to code seven studies of different research orientations and resolved all disagreements via discussion. Next, the first and second author independently coded 18 (16%) studies, resulting in high intercoder reliability (Cohen's kappa, $\kappa = .92$). The second author then proceeded to code the remaining studies. Once all studies were coded, the first author coded another two studies randomly selected from the remaining studies and the inter-rater reliability for the two studies was k=.91.

3.3. Analyses

Following prior synthesis research (Liu & Brown, 2015; Visonà & Plonsky, 2020; Zhang, Akoto, & Li, 2021), the analyses were straightforward. After studies were coded in an Excel spreadsheet, we calculated frequencies and percentages for each item of the coding scheme using SPSS. Values obtained for different categories were used to answer the RQs. It should be noted that in some tables in Section 4, the sum of the values reaches above the total number of the studies reviewed (K = 113) because a few studies apply to more than one category.

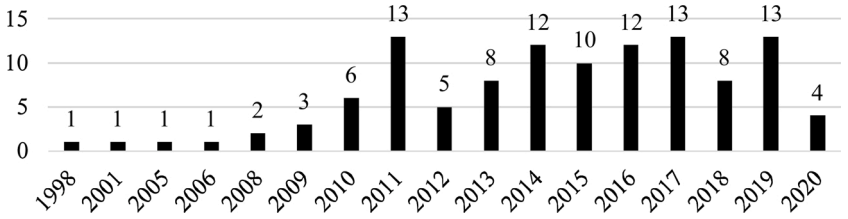

Fig. 1. Publication of Computer-mediated CW Studies Over Time.

Table 1
Sources of Computer-mediated CW Studies.

Source	*k*	*%*
Journal article	83	73
Ph.D. dissertation	11	10
Conference proceeding	8	7
M.A. thesis	6	5
Book chapter	5	4

Note. k = subset of the sample; total sample (K) = 113.

Table 2
Source Journals for Studies of CW.

Journal	*k*	*%*
ComputerAssisted Language Learning	13	12
Language Teaching and Technology	9	8
CALICO	5	4
System	4	4
ReCALL	3	3
Journal of Second Language Writing	3	3
Modern Language Journal	2	2
Australian Journal of Educational Technology	2	2
European Journal of Open, Distance, and E-learning	2	2
Interactive Learning Environments	2	2
Multimedia-Assisted Language Learning	2	2
The Turkish Online Journal of Educational Technology	2	2
Other journals[*]	34	30

[*] one article from each of these journals: *Asian Social Science; Assessment & Evaluation in Higher Education; Australian Review of Applied linguistics; Classroom Discourse; Colombian Applied Linguistics Journal; Computers in Human Behavior; Education and Information Technologies; Education Technology Research & Development; Educational Technology & Society; Electronic Journal of Foreign Language Teaching; English Language Teaching; English for Specific Purposes; Eurasian Journal of Applied Linguistics; Foreign Language Annals; IAFOR Journal of Language Learning; Innovation in Language Learning and Teaching; International Education Studies; International Journal of Computer-Assisted Language Learning & Teaching; International Journal of Computer-Supported Collaborative Learning; International Journal of English Language Education; International Journal of Modern Education and Computer Science; International Journal on E-Learning Practices; Journal of College Teaching & Learning; Journal of Information Technology Research; Language Learning; Language Education and Assessment; Language, Linguistics, Literature; Procedia-Social and Behavioral Sciences; Technology, Pedagogy and Education; The Asian Journal of English Language & Pedagogy; The International Association for Language Learning Technology Journal; The Journal of Asia TEFL; The Malaysian Online Journal of Educational Technology; Turkish Online Journal of Distance Education*.

Table 3
Coding Scheme.

Category	Items
Study ID	author(s); year; type; journal/conference title
Study Context	sample size; educational setting (foreign, L2, heritage, immersion); institution (elementary, middle school, high school, post-secondary); proficiency; age; L1; target language
Study Design	theoretical framework; research methodology (qualitative, quantitative, mixed); Qualitative research approaches (e.g., case study, grounded theory, action research); Quantitative research approaches (e.g., quasi-experimental, survey, descriptive); *independent variable(s)*[*]*; dependent variable(s)*
Study Setup	length of study; group size; grouping method; task setting; technology; mode of collaboration (synchronous, asynchronous, hybrid); # of CW tasks; task type; CW tasks
Analyses	**Dichotomous items**[**]: written products; chats/discussion; revisions; LREs; dynamics of collaboration; perception **Non-dichotomous items:** lexical complexity measure; syntactic complexity measure; accuracy measure; fluency measure; text quality measurement; coherence measurement; analysis of chats; analysis of revisions; evaluation of perception
Research & reporting practices	**Non-dichotomous items:** validity strategies for qualitative research; *statistical analysis*; *reporting of p values; effect size* **Dichotomous items:** prior training in CW; prior training in the technology; homemade prompt; piloting prompt; piloting text quality rubric; piloting survey; piloting pre-test; piloting post-test; reporting reliabilities (for coding revisions, LREs, dynamics of collaboration, accuracy, complexity, text quality, and survey); *assumption checking*; *raw frequency; percentage; mean; standard deviation; confidence interval*

[*] Note. Italicized items were only applicable to studies employing quantitative methods.
[**] Dichotomous items were coded as 0 = absent, 1 = present.

4. Results

4.1. Results for RQ 1: research contexts, demographics, and study setup

To answer RQ1, this section reports the research contexts, learner profile, and multiple study setup features (e.g., grouping method, writing tasks, technologies) of the 113 studies. As shown in Table 4, the majority of the studies were conducted in foreign language settings (64%) and post-secondary contexts (81%). Also, only six target languages have been investigated, with English accounting for 85% of the studies. Table 5 shows that on average, studies in this domain used 2.34 tasks and have a mean sample size of 41 with a large variance ($SD = 45.39$).

Concerning learner demographics, this line of research favored adult learners (81%) and intermediate proficiency (36%). The most frequently examined linguistic backgrounds were mixed L1s, Chinese, and Arabic. Also noteworthy was the missing data in reporting proficiency and L1 background. Forty-two percent of the studies did not state the learners' proficiency level or provide learners' scores in recognized English achievement tests (e.g., TOEFL, IELTS, TOEIC) which were convertible into proficiency levels based on established criteria. Instead, these studies either described learners' prior L2 learning experience (e.g., number of years learning the L2) or provided no information on proficiency. Additionally, 48% of the data did not report learners' L1 backgrounds although they often specified the country or region where the study was conducted (Table 6).

In setting up the research, 39% of the studies *assigned* learners into dyads or groups, as shown in Table 7. As to the length of studies, the majority lasted 5-9 weeks or over 10 weeks, at 33% and 35% of the studies respectively. Meanwhile, 22% reported no timeframe. Also, small groups received more attention than dyads and relatively larger groups, a pattern differing from F2F CW research which predominantly investigated dyads (Zhang & Plonsky, 2020). Moreover, 32% and 12% of studies respectively did not report how pairs/groups were formed or the group size.

Tables 8 and 9 and Figs. 2 and 3 presented information about the tasks and the technologies. As revealed in Table 8, CMCW research often implemented the tasks in non-classroom settings (k = 75, 66%). Additionally, 94% of the studies examined meaning-focused tasks, and only one study employed form-focused tasks. This large imbalance was not shared by F2F CW studies, as Zhang and Plonsky (2020) reported that 30.9% of their sample examined form-focused tasks. Also noteworthy was that 19% of the studies did not report the task settings.

Regarding the specific tasks presented in Fig. 2, we found that this scope of inquiry employed a broad range of tasks, among which expository essay, argumentative writing, and narrative/story writing were the most frequently employed. Also, tasks such as poster/brochure design and graphic writing, which were rarely examined in other areas of writing research (Johnson, 2017; Liu & Brown, 2015), have gained interests from CMCW researchers.

Regarding the task modality, Table 9 shows that all three modes (i.e., *synchronous*, real-time collaboration; *asynchronous,* non-real-time collaboration*; hybrid*) were all well-represented in this domain, with asynchronous modality receiving moderately more attention than others. Related to this, Fig. 3 which indicates that while the technologies employed in the 113 studies ranged from digital writing software (e.g., Etherpad), social networking sites (e.g., Facebook), and videoconferencing software (e.g., Skype), researchers predominantly relied on two tools: wikis and Google Docs. The predominance of the wiki, a technology noted for asynchronous communication, may account for the relatively more attention to asynchronous modality in this domain.

4.2. Results for RQ 2: Methodological design features

RQ 2 asks about methodological design features in this domain of inquiry. A number of salient patterns can be observed in Table 10. First, a qualitative methodology was much more commonly adopted than a quantitative methodology in this domain. In particular,

Table 4
Research Contexts in CMCW Research.

Context	k	%
Educational Setting		
Second language	38	34
Foreign language	72	64
Not reported	3	3
Institution		
Post-secondary	92	81
High school	10	9
Middle school	6	5
Elementary school	3	3
Not reported	2	2
Target language		
English	96	85
Spanish	7	6
German	7	6
French	2	2
Chinese	2	2
Greek	1	1
Not reported	1	1

Table 5
Sample Size and Number of Tasks.

Feature	Mean	SD	Min-max
Sample size	41.01	45.39	3 - 401
CW tasks	2.34	1.59	1–54

Table 6
Learner Demographics in CMCW Research.

Learner information	k	%
Age		
Adults	91	81
Teens	12	11
Children	8	7
Not reported	2	2
Proficiency		
Advanced	16	14
Mid-advanced	21	19
Intermediate	41	36
Low-intermediate	9	8
Low	5	4
Not reported	47	42
L1 background		
Mixed L1s	24	21
Chinese	14	12
Arabic	6	5
Turkish	4	4
Spanish	3	3
Korean	2	2
English	2	2
Japanese	2	2
Persian	1	1
Thai	1	1
Not reported	54	48

Table 7
Length of Study, Grouping, and Group Size.

Study Setup	*k*	%
Length of Study		
1–4 weeks	11	10
5–9 weeks	38	33
10 weeks or more	39	35
Not reported	25	22
Grouping method		
Self-selected	29	26
Assigned	44	39
Both	4	4
Not reported	36	32
Group size		
Pair	30	27
Group of 3–5	62	55
Group of 6 or more	14	12
Not reported	13	12

only 30 studies (27%) employed a quasi-experimental design, suggesting limited interest in this methodology by CMCW researchers. By contrast, the case study approach was implemented in 67 studies (59%) of the studies. Grounded theory and action research were sparsely employed. Also, 28% of our sample incorporated mixed methodologies in their investigations.

To identify areas in need of more future experimental research, we analyzed the independent and dependent variables in the 30 quasi-experimental studies. As revealed in Fig. 4, the top three most investigated independent variables were CW in a CM mode vs. a face-to-face mode (k = 11, 37%), CM CW versus individual writing (k = 9, 30%), and CMCW as a treatment (k = 7, 23%). Only a few studies examined variables relating to task design or learner (e.g., task type, proficiency). That is, the majority of these studies were devoted to examining the utility of CMCW by probing it as a treatment or comparing it with related tasks or modes. Regarding the dependent variables shown in Fig. 5, research in this area favored the quality of the written products (k = 15, 50%), writing gains (k = 11, 37%), and perception of the technology/task (k = 10, 33%), denoting much attention to affective aspects and more interest in the

Table 8
Task Setting and Type.

Task Feature	k	%
Task setting		
Classroom	49	43
Non-classroom	75	66
Not reported	21	19
Task type		
Meaning-focused	106	94
Form-focused	1	1
Not reported	6	6

Table 9
Mode of Collaboration in CMCW Research.

Mode of Collaboration	*k*	*%*
Synchronous	33	29
Asynchronous	46	41
Hybrid	34	30

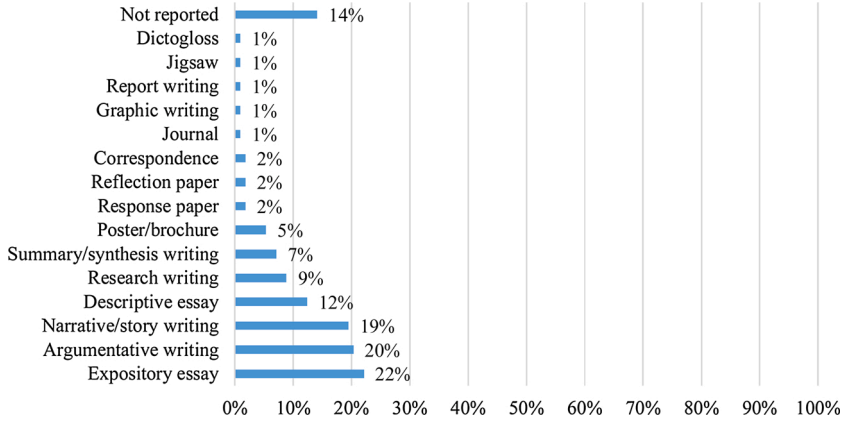

Fig. 2. Tasks in CMCW Research.

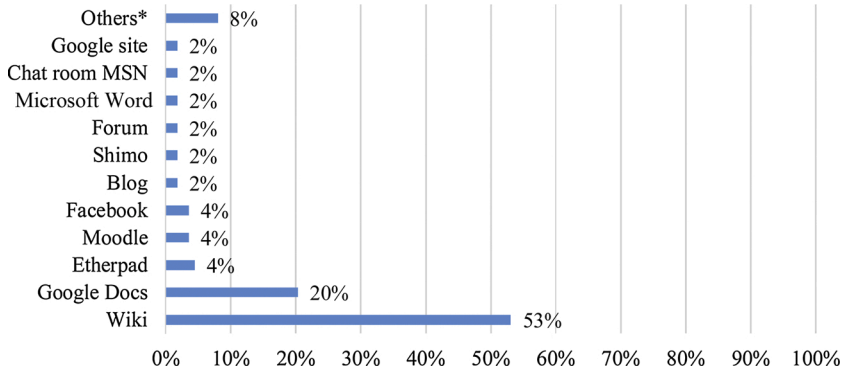

Fig. 3. Technologies for CMCW.
Note: * Other platforms include Skype, Quip, Telegram, Comic Life 3, Storybird, Systeme-D, MySpace, Padlet, Process-Writing Wizard, and WRITeam.

Table 10
Research Methodologies in CMCW Research.

Methodology	k	%
Research orientation		
Qualitative only	51	45
Quantitative only	30	27
Mixed-methods	32	28
Qualitative approach*		
Case study	67	59
Grounded theory	1	1
Action research	9	8
Quantitative approach**		
Quasi-experimental	30	27
Descriptive	28	25
Survey	4	4

* Note: the numbers under this category were based on 51 qualitative studies and 26 mixed-method studies; the remaining six mixed-method studies did not employ a specific qualitative approach but elicited qualitative data from interviews.
** the numbers under this category were based on 30 quantitative studies and 32 mixed-method studies.

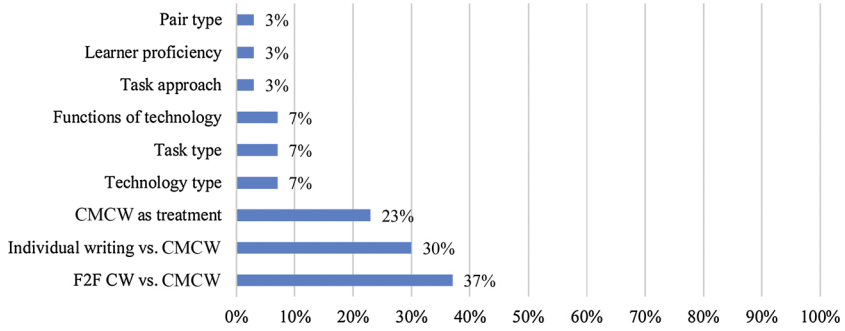

Fig. 4. Independent Variables in Quasi-experimental Studies (k = 30).

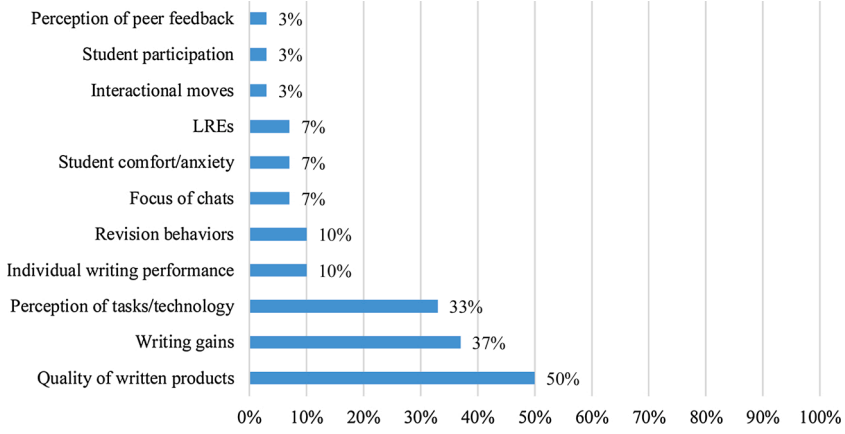

Fig. 5. Dependent Variables in Quasi-Experimental Studies (k = 30).

products than the process. A further point of interest involved the finding that only one or a few studies examined variables concerning peer interaction (e.g., LREs, the focus of chats, interactional moves). That is, limited attention was paid to peer interaction in these studies, a tendency diverging from quasi-experimental research on F2F CW.

4.3. Results for RQ3: theoretical frameworks and aspects analyzed

RQ3 addresses two related themes: (a) theoretical frameworks and (b) what and how aspects of CMCW have been analyzed. A

Table 11
Theoretical Frameworks in CMCW Research.

Theoretical framework	k	%
Constructivism/sociocultural theory	80	71
Interactionist approach	7	6
Process writing approach	3	3
Other*	8	7
None reported	28	25

* *Note*: other frameworks include social presence theory, conversation analysis framework, systemic functional grammar, dual-coding theory, project-based learning theory, politeness theory, framework of autonomous, ecological perspective.

salient feature revealed in Table 11 was that 71% of the included studies employed constructivism/sociocultural theory (Vygotsky, 1978), whereas only 6% of the studies drew on an interactionist approach (Long, 1980). Overall, although 25% of the included studies did not identify with a theoretical framework, CMCW researcher drew upon a wide array of additional theories, ranging from theories in linguistics/applied linguistics (e.g., politeness theory by Brown, Levinson, & Levinson, 1987; autonomous learning theory by Littlewood, 1996; systemic functional grammar by Halliday, 1973; the ecological perspective by Van Lier, 2004) to theories in social psychology and cognition (e.g., dual-coding theory by Paivio, 1971; social presence theory by Short, Williams, & Christie, 1976).

With regard to the aspects that were analyzed in this domain, 67% of the studies addressed learners' perception of the technology or the task, suggesting much attention to learners' attitudinal evaluation. Also, much effort was devoted to examining learner interaction in the form of chats or online discussion (k = 54, 48%), the co-constructed texts (k = 51, 45%), and the revising processes (k = 43, 38%). In contrast, LREs have been examined in only 23% of the studies (Fig. 6).

Aside from answering *what* has been analyzed in this domain, we were also interested in *how* each aspect was analyzed. The findings were displayed in Figs. 7–9 and Tables 12–14. As illustrated in Fig. 7, learners' online chats/discussion in CMCW tasks were primarily analyzed to shed light on the focus of the interaction (k = 41, 36%) and the level of engagement (k = 31, 27%), Occasionally, researchers directed their attention to examining the LREs, the pragmatic function, and how teachers intervened the peer interaction.

Table 12 reports how revisions and task/technology perception were examined. In analyzing revisions, 38% of the studies examined the different types of changes made to the texts and 5% of the studies investigated revisions as a means to discern the amount or the type of contribution that each pair/group member made to the collaborative task. In examining learners' perception of the technology and/or the CMCW task, studies in this area utilized diverse methods, among which survey and interview were the most widely employed. Also worth noting was the finding that 32% of the studies used two or more methods to gauge learners' perception, which was indicative of triangulation in this regard.

Regarding the analyses of co-constructed texts, the included studies heavily relied on the construct of text quality and the CAF framework, as demonstrated in the top five aspects analyzed in Fig. 8. Nevertheless, a small portion of studies probed into less traditional constructs such as coherence/cohesion (k = 4, 4%), creativity (k = 1, 1%), and textuality (k = 1, 1%).

When zooming in on the specific measures in Table 13, we found that 16 different syntactic measures were used in the 16 studies that examined syntactic complexity. This heterogeneity was in line with research on task complexity and task-based production (Johnson, 2017; Plonsky & Kim, 2016). They fell under three categories: length-related, subordination-related, and other measures. Except for the mean length of the T-unit and clauses per T-unit which were respectively employed in 5% and 4% of the studies, the other measures were occasionally used across studies. Also, one study employed noun modifiers as a measure of syntactic complexity. With regard to the lexical aspects, the type token ratio was most frequently employed (k = 9, 8%).

The heterogeneity in measurement holds for linguistic accuracy and fluency as well. As presented in Fig. 9, 12 different measures were much evenly employed in these studies. Table 14 presents eight measures for linguistic fluency, with text length being the most widely employed. Further, coherence/cohesion was evaluated through either a holistic rubric, coherence breaks, or selective linguistic features. Text quality was rated holistically twice more often than analytically in the included studies.

4.4. Results for RQ4: research and reporting practices

In this section, we first presented findings on research and reporting practices that were not necessarily associated with a particular research methodology: (a) instrument piloting, (b) pre-task training, and (c) the reporting of reliability. We then respectively presented

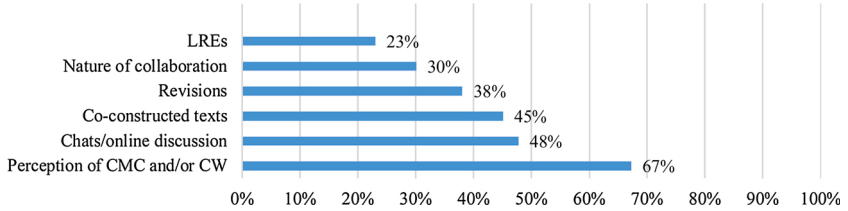

Fig. 6. Features Analyzed in CMCW Research.

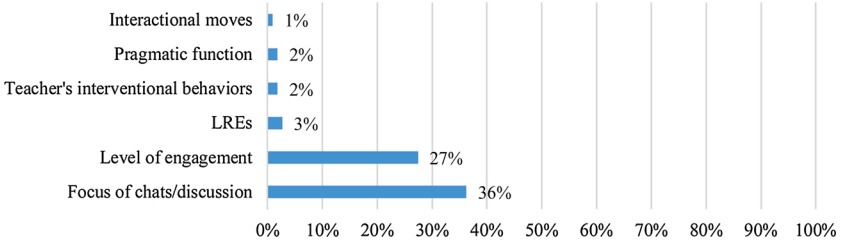

Fig. 7. Analysis of Chats and Online Discussion.

Table 12
Analyses of Revisions and Perception.

Analysis	k	%
Analysis of revisions		
Type of revisions	38	34
Task contribution	5	4
Perception of CMC and/or CW		
Survey	63	56
Interview	42	37
Journal/reflective paper	12	11
Observation notes	8	7
2 or more abovementioned methods	36	32

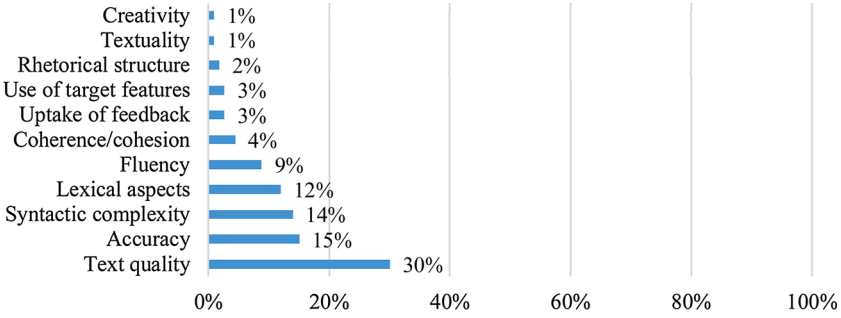

Fig. 8. Analyses of Co-constructed Texts.

methodological findings that were applicable to studies adopting quantitative methods and those applicable to studies employing qualitative methods.

As appeared in Fig. 10, 56% of the included studies reported the implementation of pre-task training in the technology being used, whereas only 32% of the studies reported conducting pre-task training in the CW approach. In terms of instrument development and piloting, 46% of studies employed homemade writing prompts, and 11% reported piloting the prompt. Also, pre- and post-tests were reported to be piloted in only 27% of the studies.

To more accurately document researchers' efforts in reporting reliability, we distinguished three situations: reliability coefficient reported, not reported, and not reported yet the coding reported to be verified by an additional coder. The results in Fig. 11 indicated that reliability was most frequently reported in the coding of text quality (68%), coherence/cohesion (50%), and linguistic accuracy (47%). Among the low rates of reporting reliability for various features, the most striking was the coding of revisions and interviews. While occasionally researchers reported having a second coder check the coding, only 14% of the studies reported reliability for coding interview data and only 23% reported that for coding revisions.

Following prior reviews of quantitative research (e.g., Plonsky, 2013; Plonsky & Kim, 2016), we examined the inferential statistics and features associated with researcher transparency for the studies that employed any quantitative method in our sample (i.e., quantitative and mixed-method studies). Fig. 12 indicated that among the 12 different inferential statistics employed, the most commonly used statistical techniques were the *t*-test (k = 24, 39%), ANOVA (k = 8, 13%), Mann-Whitney *U* test (k = 7, 11%), and Wilcox signed-ranked test (k = 7, 11%). Very few studies used multiple regression or generalized estimating equations.

Regarding the eight features associated with researcher transparency, Fig. 13 illustrated that while the mean, raw frequency, and standard deviation were reported more often, four features were reported only occasionally, including exact *p* values for all tests (k = 13, 30%), assumption checking (k = 12, 28%), confidence interval (k = 11, 18%), and effect sizes for all tests (k = 3, 7%).

Since the above-mentioned items were not relevant to qualitative research, we examined the validity strategies in studies that

Table 13
Measures of Syntactic Complexity and Lexical Aspects.

Measures	*k*	%
Syntactic complexity	16	14
Length related measures		
Mean length of T-unit	6	5
Mean length of clause	2	2
Mean length of sentence	2	2
Mean length of c-unit	1	1
Mean period unique length	1	1
Subordination related measures		
Clauses per T-unit	4	4
Dependent clauses/total clauses	2	2
Subordinate clauses/total clauses	1	1
Subordinate clauses per 100 words	1	1
Clauses per c-unit	1	1
Subordination and coordination per sentence	1	1
Cohesive conjunctions per sentence	1	1
Other measures		
Number of modifiers per noun phrase	1	1
Number of T-unit	1	1
Unique bigram ratio	1	1
Measures of lexical aspects	14	12
Type token ratio	9	8
% of words in academic word list	1	1
Mean word length	1	1
Fog Index	1	1
% of lexical words	1	1
% of 100 most frequent words	1	1

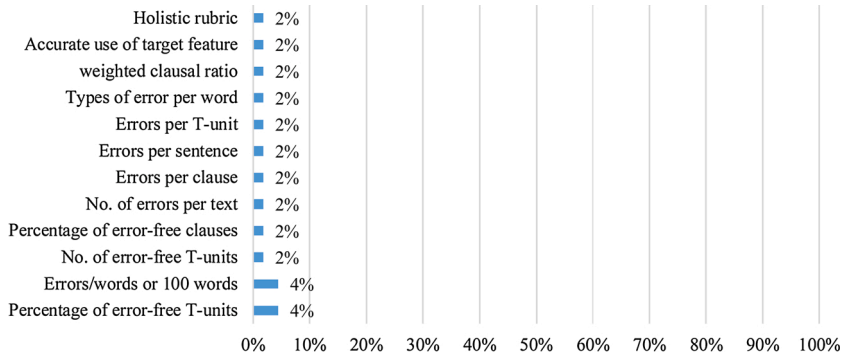

Fig. 9. Accuracy Measures in CMCW Research.

Table 14
Measurement of Fluency, Coherence/Cohesion, and Text quality.

Measures	*k*	%
Fluency		
Text length	8	7
No. of T-unit	5	4
No. of clauses	2	2
Mean length of clause	2	2
Mean length of T-unit	2	2
Words per minute	2	2
No. of sentences	1	1
Dependent clauses/total clauses	1	1
Coherence/cohesion		
Holistic rating	2	2
Coherence breaks	1	1
Selective linguistic features	1	1
Text quality		
Holistic rating	23	20
Analytic rating	11	10

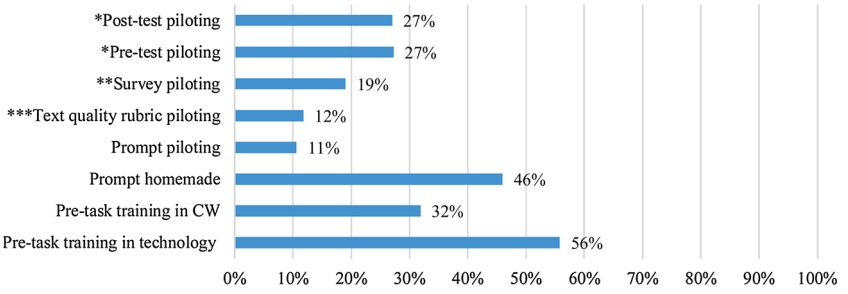

Fig. 10. Instrument Development and Pre-task Training.
Note: * percentage based on 22 studies that employed a pre- and post-test designs.
**percentage based on 63 studies that used surveys.
***percentage based on 34 studies that used text quality rubrics.

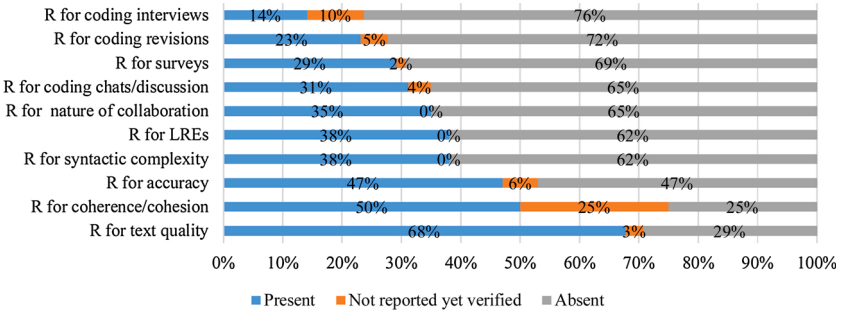

Fig. 11. Reporting of Reliability in CMCW Research.

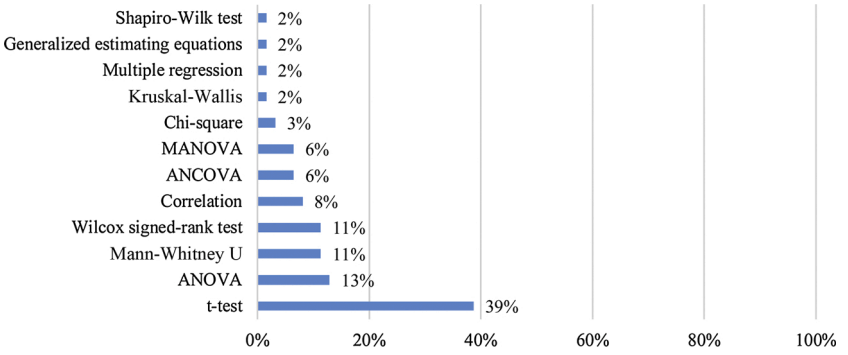

Fig. 12. Inferential Statistics in Quantitative and Mixed-method Studies (k = 62).

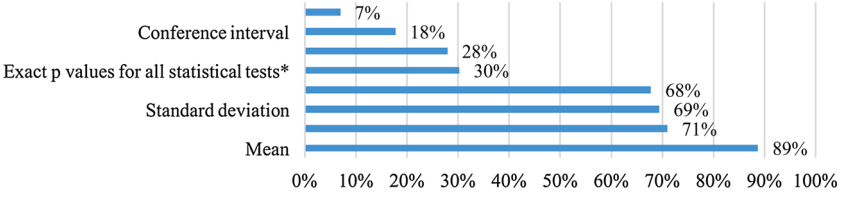

Fig. 13. Features Associated with Researcher Transparency in Quantitative and Mixed-method Studies (k = 62).
**Note:* these items were only applicable to the 43 studies that employed inferential statistics.

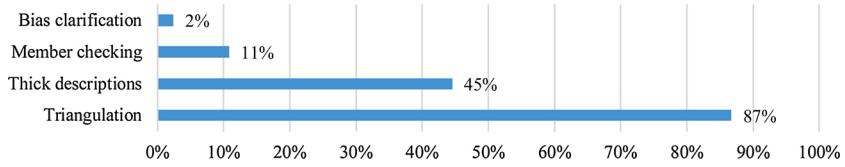

Fig. 14. Validity Strategies in Qualitative and Mixed-method Studies (k = 83).

employed qualitative methods (i.e., qualitative and mixed-method studies) to gain insight into their research and reporting practices. As displayed in Fig. 14, the most frequently employed validity strategies were triangulation (k = 72, 87%) and thick descriptions (k = 37, 45%). By contrast, member checking was reported in only 11% of the sample, and bias clarification was reported in only 2% of the studies.

5. Discussion

This study examined CMCW research from both the substantive and methodological perspectives. We were interested not only in describing the contexts, learner populations, constructs, and features that had been studied in this area, but also in reviewing the methods and practices that had been employed to investigate them. In this section, we discuss our findings in relation to related areas of L2 research and consider the link between the substantive aspects and the methodological choices. Drawing upon our findings, we identify areas in need of more research and provide a number of empirically grounded suggestions for advancing this domain of research.

5.1. Sampling, contexts, and study setup

One salient pattern in our dataset was the predominant focus on adult learners, post-secondary settings, and English learners. This trend echoes multiple other L2 research domains such as task-based research (Plonsky & Kim, 2016), computer-assisted instruction (Sharifi, Rostami AbuSaeedi, Jafarigohar, & Zandi, 2018), and written corrective feedback (Liu & Brown, 2015). While prevalent in L2 research, the underrepresentation of non-adult learners and learners of other languages are problematic (Andringa & Godfroid, 2020; Ortega, 2005) in that it limits the generalizability, leads to potential underestimation of variance in elicited learner data, and makes it difficult to provide answers to L2 practitioners who work with non-English learners, and learners of other age groups, including K-12 students.

Additionally, unlike F2F CW tasks which were often implemented in dyads and in classroom settings (Zhang & Plonsky, 2020), CMCW was more commonly conducted outside the classrooms and in small groups. This is related to the fact that technology affords opportunities for multiple learners to collaborate on a task outside the classrooms synchronously or asynchronously via a multitude of tools. The most frequently employed technologies (e.g., wiki, Google Docs) are found to be similar to those employed in collaborative learning in general (Su & Zou, 2020). Likewise, due to the affordance of technology, researchers in this area are able to investigate tasks that often require multi-modal components such as poster design (Alghasab & Handley, 2017), graphic writing (Kılıçkaya, 2020), and tourism brochure design (Diez-Bedmar & Perez, 2012). New and emerging CMC technologies provide L2 students with opportunities to develop digital literacies together with collaborative writing skills (Li & Storch, 2017, Li & Zhang, 2021). Nevertheless, research in this domain shared interests in expository, argumentative, and narrative writing tasks with F2F CW research (Zhang & Plonsky, 2020). This may be related to the sample's heavy focus on adult, post-secondary, and intermediate to advanced learners who are often exposed to such tasks. Meanwhile, the scant attention to form-focused tasks, a pattern diverging from F2F CW research, may suggest that CMCW tasks are deemed as a site primarily for meaning co-construction in the L2.

5.2. Research methodologies in computer-mediated CW research

With regard to the research methodologies, qualitative methods, especially the case study approach, are indispensable to this scope of inquiry. This preference is likely related to two factors: (a) unlike F2F CW which is often completed in class, CMCW tasks in the asynchronous or hybrid modes often last for a span of time, which allows researchers to study behaviors over a period of time; (b) crucial constructs such as online revision behaviors, acts of co-construction, and scaffolding strategies are subject to qualitative interpretations (e.g., Li & Kim, 2016; Li & Zhu, 2017a). When examining more closely, we find that mixed-method studies often involve using numeric data to describe performance in tests or writing tasks and employing qualitative methods for probing peer interaction or attitudinal evaluation (e.g., Bikowski & Vithanage, 2016). Action research is not uncommon as a method for L2 practitioners to investigate pedagogical frameworks for implementing CW online (e.g., Levrai & Bolster, 2019) or students' attitudes towards a technology or task (e.g., Chao & Lo, 2011). Grounded theory is used in only one study (Blin & Appel, 2011) to theoretically construe the mediational and collaborative structures of computer-supported CW activities via cultural-historical activity theory.

The finding that a large portion of the quasi-experimental studies is devoted to comparing CMCW with similar tasks (e.g., F2F CW, individual writing) was not surprising given that this domain is a young research area and such research is necessary to discover the utility and effectiveness of this writing approach in L2 classrooms. Building upon these studies, a small portion of studies have started to explore how task variables (e.g., task type, task approach) and learner variables (e.g., proficiency) may influence the potential of CMCW (e.g., Yeh, Lo, & Huang, 2011).

5.3. Theoretical frameworks, features, and measurement

With regard to the theoretical frameworks, 75% of the reviewed studies discuss the theoretical position and/or specific theoretical constructs that informed research on CMCW. The research favors the sociocultural theory perspective, including the notions of Zone of Proximal Development, scaffolding, and collective scaffolding. More recently, like the research area of peer response (e.g., Zhu & Mitchell, 2012), activity theory begins to capture researchers' interests in understanding the socio-cultural factors (e.g., goals, motives, emotions) that mediated dynamics of peer interaction (e.g., Li & Zhu, 2017a). Different from research on F2F collaborative writing, only a small portion of the research is undergirded by the interactionist approach. In the interactional processes regarding CMCW tasks, learners can focus on linguistic forms and notice the gap between the interlanguage and the target language (Torres & Cung, 2019). One possible reason why only a few studies draw on the interactionist approach is that the CMCW tasks focus more on content development than language forms, and learners' attention to form is not easy to capture particularly in the asynchronous CW context. Moreover, the process writing approach used in a few studies is related to the affordances of web 2.0 tools for multiple rounds of revisions and staged scaffolding. In addition, theoretical frameworks in other disciplines (e.g., politeness theory, autonomous learning, social presence theory) are found to have enlightened work on CMCW, closely related to research purposes and research tasks.

Regarding the components analyzed in our sample, two interesting patterns stand out. First, learners' revision processes, a feature rarely examined in F2F CW research, gained considerable interest in this area. This is partly because technological tools for digital writing (e.g., wikis, Google Docs) permit efficient storage of revisions made by multiple individual learners. The accessibility of such data enables researchers to gain insight into both learners revising processes and task contribution. The other striking pattern concerns LREs. While both areas are heavily informed by sociocultural theories and their related concepts (e.g., collaborative dialogue, *languaging*) (Suzuki & Storch, 2020), CMCW devoted scarce attention to LREs, which is a feature indicative of L2 learning opportunities, whereas Zhang and Plonsky (2020) reported that LREs was the most frequently examined feature in F2F CW. Two factors may have led to this discrepancy. First, as previously mentioned, the vast majority of the tasks employed in this area is meaning-focused writing tasks which typically generate less attention to form (Storch, 2013). Additionally, by affording asynchronous communication, co-composing, and co-editing at one site, technological tools alter the forms of collaborative dialogue. In CMC settings, learners' attention to form might be more 'slippery' and more difficult to capture, as they may occur in the chats and online discussion or be directly incorporated into the co-constructed texts. Researchers (e.g., Kessler, 2009; Li, 2020) have made efforts to capture less explicit attention to form by scrutinizing the revision histories for evidence of LREs. For instance, Li (2020) reported a large portion of LREs at the level of limited engagement in a study on wiki-based collaborative writing. Yet, the field could benefit from more efforts in developing methodologies for capturing learners' attention to form across different sites in order to better understand language learning opportunities afforded by CMCW tasks.

In analyzing the co-constructed texts, research in this domain widely employs the construct of text quality and the CAF framework. There are two salient traits: (a) the heterogeneity in the metrics of CAF and (b) the overlapping between certain complexity and fluency measures (see Tables 13–14 and Fig. 7). The heterogeneity is shared by research on task complexity, task-based production, and L2 writing in general (Johnson, 2017; Lan, Liu, & Staples, 2019; Plonsky & Kim, 2016;). The overlapping of complexity and fluency metrics was shared by F2F CW research (Zhang & Plonsky, 2020). Nevertheless, more consistency and improved clarity in measuring different constructs are warranted to make findings across studies more comparable and further our understanding of the written products in CMCW research.

Also noteworthy is the fact that recent research in this area is starting to pay attention to phrasal complexity and coherence/cohesion. Recent corpus investigations have consistently demonstrated the importance of phrasal complexity for measuring written texts in academic settings, as phrasal features (e.g., constituents of noun phrases) are critical devices for writers to compress rich information into concise discourse (e.g., Biber, Gray, & Poonpon, 2011; Biber, Gray, Staples, & Egbert, 2020; Parkinson & Musgrave, 2014; Zhang, 2021). It is applaudable that researchers (e.g., Yim, 2017) in this area have started to incorporate these new findings into measuring learner texts. In addition, the act of co-constructing a text by two or multiple learners in online modalities may introduce potential discontinuity in the discourse at both the micro- and macro levels, which necessitates the measurement of coherence and cohesion, essential aspects of writing, in CMCW research (Li & Zhu, 2017b). Future CMCW studies should consider these less explored constructs when relevant to the tasks being implemented.

5.4. Research and reporting practices

With regards to research and reporting practices, we found high percentages of studies that failed to provide crucial information about learners and the study setup, including learners' L1s, learner proficiency, grouping method, length of study, task setting, and the tasks implemented. The lack of such information is concerning in two ways. First, it undermines the interpretability of the research findings as it is difficult to consider their results in relation to studies on a similar learner population or implemented in a similar fashion. Further, it is difficult for consumers of research to contextualize the findings of these studies. Prior research has indicated that study setup features such as group size (Fernández Dobao, 2012), grouping information (Mozaffari, 2017), and tasks (Torres & Cung,

2019) may influence the collaboration process and products, which points to the need for future research to more consistently report such information.

Also, over half of the studies reported the implementation of pre-task training on the technology in use, but less than one-third reported having pre-task training in CW. While pre-task training might not be an issue in classrooms where writing collaboratively with technology is a common practice, it may pose issues in other classrooms, as previous research shows that learners' familiarity with the CW approach influences the different aspects of CW (Kim & McDonough, 2011). Moreover, like F2F CW research, CMCW also widely employs homemade prompts, which is typical of classroom-based research. Concerning instrument piloting, only six out of the 22 studies (27%) piloted the pre- and post-tests. While the absence of such information does not necessarily equal to no piloting and may relate to the word limits of journals, it remains critical to report such information to avoid potential threats to the internal validity of the instruments. Otherwise, it is difficult to discern whether the detected effects or the lack of them might have been confounded by certain design features of the pre- and post-tests.

Moreover, we found low rates of reporting the reliability of surveys and the manual coding of most L2 features (e.g., accuracy, syntactic complexity, LREs, the focus of chats/discussion). Of particular note, in coding interview data, 76% of the time researchers did not report reliability estimates or report having another rater cross-check the coding. Likewise, in coding revising behaviors, reliability or cross-checking was not reported 72% of the time. Although CMCW research is not unique in this practice and the absence of reliability does not necessarily imply unreliable coding or surveys, unreported reliability or the lack of cross-checking is problematic in multiple ways. First, it affects the interpretability of the results by making it difficult to evaluate whether the results may have been skewed by low reliability of the coding (Creswell & Creswell, 2017; McKay & Plonsky, 2021; Plonsky & Derrick, 2016). Also, it leaves future researchers who are interested in similar constructs little information as to which analytic frameworks (e.g., frameworks for analyzing revisions) are more reliable, which, in return, may hinder the refinement of the analytic frameworks. Moreover, unreported reliability in quasi-experimental studies, in particular, affects meta-analytic syntheses of primary research in this domain because meta-analyses rely on reliability estimates to weigh the effects of treatment (Plonsky, 2013).

When focusing on quantitative studies, we find that this domain shows a preference for statistical techniques (e.g., *t*-test, ANOVA, Mann-Whitney *U* tests) for testing mean differences between or among groups, a pattern resembling L2 research in general and F2F CW research (Plonsky, 2013; Zhang & Plonsky, 2020). Multivariate statistics are only occasionally employed in this area when sample sizes suffice. Although compared with related areas (e.g., L2 research in general, interaction research, F2F CW), this domain of inquiry seems to show a strength in reporting statistical assumption checking, standard deviations, and conference interval (e.g., Hu & Plonsky, 2021; Plonsky & Gass, 2011), the reporting of these values is far from ideal. Also, effect sizes and exact *p* values were respectively reported in merely 7% and 30% of the studies that employed inferential statistics, which are strikingly lower than research on corrective feedback and F2F CW (Liu & Brown, 2015; Zhang & Plonsky, 2020). The insufficient reporting of such information can impede the robust development of this field. For instance, the absence of standard deviation results in no information on the variance of a dataset and the potential exclusion of a study from meta-analytic reviews (Plonsky, 2014). The field can also benefit from improved reporting of effect sizes, confidence intervals, and exact *p* values, which are crucial for helping consumers of research understand the practical significance of a treatment (Larson-Hall & Plonsky, 2015).

In terms of the qualitative studies on CMCW, we found that triangulation, the use of multiple data sources and theories to provide corroborating evidence (Lincoln & Guba, 1985), is the most common strategy researchers use to validate the research. Nearly half of the reviewed studies also employed the validation strategy of providing a thick description. Future research should ensure to have the details of the participants, the setting, and the procedures described in detail so that readers can make informed decisions regarding transferability (i.e., whether the findings can be transferred because of shared characteristics). Member checking, "the most critical technique for establishing credibility" (Lincoln & Guba, 1985, p. 314), is largely missing in the current body of literature. Future qualitative research is encouraged to more often solicit participants' feedback on the accuracy of the reported content. Moreover, the reporting of researcher bias received little attention in the current literature. The researcher him/herself can be a risk to validity as the researcher is involved in the data collection process. Therefore, it is important to clarify potential researcher biases(e.g., past experiences and orientations) so that readers can understand the researcher's position and interpretation that might have impacted the inquiry (Creswell & Creswell, 2017; Richards, 2009).

5.5. *Implications of the study*

Through reviewing the empirical landscape of CW in digital settings, this study provides both theoretical and pedagogical implications. Theoretically, findings of this review suggest that CMCW has been largely construed as a site for meaning co-construction by L2 writers, which diverges from the fact that F2F CW tends to be more commonly seen as a site for both meaning and language co-construction (Li & Zhang, 2021; Zhang & Plonsky, 2020). While this pattern relates to the fact that language co-construction (e.g., LREs) can be more difficult to capture in digital settings (particularly in asynchronous or hybrid modes), embracing a broader view of CW in digital settings is to advance this sub-domain. Further, the fact that theoretical frameworks such as politeness theory and social presence theory have been employed in the selected studies illuminates additional avenues for expanding the boundaries of CW as a construct. For instance, via the lens of social presence theory, the prominent identity of those who engage in CMCW activities is no longer a L2 *learner* or *writer*, but a *communicator* in the online environment. Pedagogically, the findings suggest that CMCW tasks can be implemented in L2 classrooms in a flexible fashion to suit different instructional needs and goals: it is compatible with a wide range of genres and can be carried out in different modalities via various technological tools.

5.6. Recommendations for future research

The results of this study lead to a number of suggestions for future research. First, there is a pressing need for future research to diversify the participant sample by including non-English learners, non-adult learners, and secondary settings. Second, there is little doubt that qualitative techniques such as case study remain highly valuable for studying learners' behaviors and the interplay of variables over a period of time in a fine-grained manner. While being cognizant of the fact that interventionist research may not best align with the dominant theorectical framwork in this domain (i.e., sociocultural theory) and the vast challenges involved in conducting interventionist research in CMC settings, we argue that when suiting the research goals, experimental research that controls confounding variables are warranted to complement this domain of inquiry. Existing experimental research has devoted much attention to answering the question of *whether* CMCW is useful in L2 classrooms. Looking ahead, more experimental or mixed-method research is needed to answer the question of *how* the utility of CMCW may be influenced by variables relating to task design and implementation (e.g., group size, task type, task approach, use of writing strategies) and technology-related variables (e.g., synchronous vs. asynchronous functions, technology type). Learner variables (e.g., learner beliefs, learner identity, learner agency, task-specific motivation), which also constitute important research topics in this domain, can be researched through mixed-method studies that incorporate in-depth inquiries and quantitative measurement of learner performance or development in CW.

Further, given that existing research has mainly investigated CMCW as a task for co-constructing meaning, it may be worth exploring whether CMCW has the potential to provide a site for co-construct attention to form by examining form-focused writing tasks such as dictogloss. Another two areas that are critical yet underexplored are activities at the pre-writing stage (e.g., pre-task planning/modeling) and assessment practices that can meet the new opportunities and challenges brought about by new digital technologies (Li & Storch, 2017; Li & Zhang, 2021; Zhang & Chen, forthcoming). In terms of analyses, this sub-domain can benefit from more consistent metrics of CAF, employment of cohesion/cohesion and phrasal complexity (when relevant), and more nuanced methodologies for capturing attention to form in asynchronous and hybrid modes. Finally, we call for improved research and reporting practices regarding learner profiles (e.g., L1 background, proficiency), study setup (e.g., grouping method, group size, pre-task training), reliability estimates, statistical information for quantitative studies (e.g., standard deviation, effect sizes, exact p values), and validity strategies (especially member checking and research bias clarification) for qualitative research.

6. Conclusion

This study reviewed 113 primary studies on CMCW between January 1998 and May 2020. The findings illustrate that informed by a variety of theories, CMCW research has employed diverse methodologies (e.g., case study, experimental design, action research) to examine peer interaction, co-composing process, collaborative texts, and stakeholders' attitude towards this pedagogic task. By examining the substantive and methodological aspects of the studies, this study illuminates the existing scope of inquiry, provides an account of existing analytic frameworks, locates areas in need of future research, and identifies both strengths and limitations of existing research and report practices in this domain. Nevertheless, when considering the aforementioned suggestions, we must caution readers about one limitation of this review: although research on CMCW has been reported in other languages around the world, our review focused exclusively on work written in English, which might have influenced our findings (especially regarding the learner profile). Reviews of CMCW research published in other languages are needed to afford a fuller picture of this research area. Finally, having witnessed how technologies have empowered researchers in this domain to examine new writing genres, task modes, and composing processes in unprecedented ways (Oskoz & Elola, 2020), we are convinced that this domain of inquiry is to be continuously invigorated by advancement in technology and look forward to more primary research that furthers our understanding of CMCW in various settings.

References

Abrams, Z. (2016). Exploring collaboratively written L2 texts among first-year learners of German in Google Docs. *Computer Assisted Language Learning, 29*(8), 1259–1270. https://doi.org/10.1080/09588221.2016.1270968

Abrams, Z. (2019). Collaborative writing and text quality in Google Docs. *Language Learning & Technology, 23*(2), 22–42.

Alghasab, M., & Handley, Z. (2017). Capturing (non-) collaboration in wiki-mediated collaborative writing activities: The need to examine discussion posts and editing acts in tandem. *Computer Assisted Language Learning, 30*(7), 664–691. https://doi.org/10.1080/09588221.2017.1341928

Al-Hoorie, A. H., & Vitta, J. P. (2019). The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research, 23*, 727–744. https://doi.org/10.1177/1362168818767191

Andringa, S., & Godfroid, A. (2020). On the foundations of knowledge in applied linguistics research: Sampling bias and the problem of generalizability. *Annual Review of Applied Linguistics, 40*, 134–142. https://doi.org/10.1017/S0267190520000033

Antón, M., & DiCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review, 54*(3), 314–342. https://doi.org/10.3138/cmlr.54.3.314

Arnold, N., Ducate, L., & Kost, C. (2012). Collaboration or cooperation? Analyzing group dynamics and revision processes in wikis. *CALICO Journal, 29*(3), 431–448. https://doi.org/10.11139/cj.29.3.431-448

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

Biber, D., Gray, B., Staples, S., & Egbert, J. (2020). Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes*, 100869. https://doi.org/10.1016/j.jeap.2020.100869

Bikowski, D., & Vithanage, R. (2016). Effects of web-based collaborative writing on individual L2 writing development. *Language Learning & Technology, 20*(1), 79–99.

Blin, F., & Appel, C. (2011). Computer supported collaborative writing in practice: An activity theoretical study. *CALICO Journal, 28*(2), 473–497. https://doi.org/10.11139/cj.28.2.473-497

Brown, J. D. (2005). Research methods for applied linguistics. In A. Davies, & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 476–500). Oxford: Blackwell.

Brown, P., Levinson, S. C., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge: Cambridge university press.

Byrnes, H. (2013). Notes from the editor. *Modern Language Journal, 97*, 825–827. https://doi.org/10.1111/j.1540-4781.2013.12051.x
Chao, Y. C. J., & Lo, H. C. (2011). Students' perceptions of Wiki-based collaborative writing for learners of English as a foreign language. *Interactive Learning Environments, 19*(4), 395–411. https://doi.org/10.1080/10494820903298662
Cho, H. (2017). Synchronous web-based collaborative writing: Factors mediating interaction among second-language writers. *Journal of Second Language Writing, 36*, 37–51. https://doi.org/10.1016/j.jslw.2017.05.013
Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
Denzin, N. K. (1989). *Interpretive interactionism*. Newbury Park, CA: Sage. https://doi.org/10.4135/9781412984584
Díez-Bedmar, M. B., & Pérez-Paredes, P. (2012). The types and effects of peer native speakers' feedback on CMC. Language Learning &. *Technology, 16*(1), 62–90.
Donato, R. (1994). Collective scaffolding in second language learning. In J. Lantolf, & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.
Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
Ducate, L. C., Anderson, L. L., & Moreno, N. (2011). Wading through the world of wikis: An analysis of three wiki projects. *Foreign Language Annals, 44*(3), 495–524. https://doi.org/10.1111/j.1944-9720.2011.01144.x
Elabdali, R. (2021). Are two heads really better than one? A meta-analysis of the L2 learning benefits of collaborative writing. *Journal of Second Language Writing*, Article 100788.
Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering L2 development and mastery of writing conventions. *Language Learning & Technology, 14*(3), 51.
Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication, 32*(4), 400–414. https://doi.org/10.2307/356602
Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*, 40–58. https://doi.org/10.1016/j.jslw.2011.12.002
Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Arnold.
Heigham, J., & Croker, R. (Eds.). (2009). *Qualitative research in applied linguistics: A practical introduction*. London, England: Palgrave Macmillan. https://doi.org/10.1057/9780230239517.
Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461–473. https://doi.org/10.1093/applin/amp048
Hsu, H. C., & Lo, Y. F. (2018). Using wiki-mediated collaboration to foster L2 writing performance. *Language Learning & Technology, 22*(3), 103–123.
Hu, Y., & Plonsky, L. (In press). Statistical assumptions in L2 research: A systematic review. Second Language Research, 0267658319877433.
Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing, 37*, 13–38. https://doi.org/10.1016/j.jslw.2017.06.001
Kessler, G. (2009). Student-initiated attention to form in wiki-based collaborative writing. *Language Learning & Technology, 13*(1), 79–95.
Kessler, G., & Bikowski, D. (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. https://doi.org/10.1080/09588220903467335
Kılıçkaya, F. (2020). Learners' perceptions of collaborative digital graphic writing based on semantic mapping. *Computer Assisted Language Learning, 33*(1-2), 58–84. https://doi.org/10.1080/09588221.2018.1544912
Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal, 92*(1), 114–130.
Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research, 15*(2), 183–199. https://doi.org/10.1177/1362168810388711
Kost, C. (2011). Investigating writing strategies and revision behavior in collaborative wiki projects. *CALICO Journal, 28*(3), 606–620. https://doi.org/10.11139/cj.28.3.606-620
Lai, C., Lei, C., & Liu, Y. (2016). The nature of collaboration and perceived learning in wiki-based collaborative writing. *Australasian Journal of Educational Technology, 32*(3). https://doi.org/10.14742/ajet.2586
Lan, G., Liu, Q., & Staples, S. (2019). Grammatical complexity: 'What Does It Mean' and 'So What' for L2 writing classrooms? *Journal of Second Language Writing, 46*. https://doi.org/10.1016/j.jslw.2019.100673
Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(S1), 127–159. https://doi.org/10.1111/lang.12115
Lei, L., & Liu, D. (2019). Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics, 40*(3), 540–561. https://doi.org/10.1093/applin/amy003
Levrai, P., & Bolster, A. (2019). A framework to support group essay writing in English for Academic Purposes: A case study from an English-medium instruction context. *Assessment and Evaluation in Higher Education, 44*(2), 186–202. https://doi.org/10.1080/02602938.2018.1487024
Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning, 31*(8), 882–904. https://doi.org/10.1080/09588221.2018.1465981
Li, M. (2020). Languaging in wiki-based collaborative writing: Functions and mediating factors. In N. Storch, & S. Wataru (Eds.), *Languaging in language teaching and learning: A collection of empirical studies* (pp. 149–169). Philadelphia, PA: John Benjamins.
Li, M., & Kim, D. (2016). One wiki, two groups: Dynamic interactions across ESL collaborative writing tasks. *Journal of Second Language Writing, 31*, 25–42. https://doi.org/10.1016/j.jslw.2016.01.002
Li, M., & Storch, N. (2017). Second language writing in the age of CMC: Affordances, multimodality, and collaboration. *Journal of Second Language Writing, 36*, 1–5. https://doi.org/10.1016/j.jslw.2017.05.012
Li, M., & Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Computer Assisted Language Learning, 26*(1), 62–81. https://doi.org/10.1080/09588221.2011.631142
Li, M., & Zhu, W. (2017a). Explaining dynamic interactions in wiki-based collaborative writing. *Language Learning & Technology, 21*(2), 96–120.
Li, M., & Zhu, W. (2017b). Good or bad collaborative wiki writing: Exploring links between group interactions and writing products. *Journal of Second Language Writing, 35*, 38–53. https://doi.org/10.1016/j.jslw.2017.01.003
Li, M., & Zhang, M. (2021). Collaborative writing in L2 classrooms: A research agenda. *Language Teaching*. https://doi.org/10.1017/S0261444821000318
Lincoln, Y. G., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
Littlewood, W. (1996). "Autonomy": An anatomy and a framework. *System, 24*(4), 427–435. https://doi.org/10.1016/S0346-251X(96)00039-5
Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30*, 66–81. https://doi.org/10.1016/j.jslw.2015.08.011
Long, M. H. (1983). Linguistic and conversational adjustments to nonnative speakers. *Studies in Second Language Acquisition, 5*, 177–193. https://doi.org/10.1017/S0272263100004848
Luppicini, R. (2007). Review of computer mediated communication research for education. *Instructional Science, 35*(2), 141–185. https://doi.org/10.1007/s11251-006-9001-6
Mak, B., & Coniam, D. (2008). Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System, 36*(3), 437–455. https://doi.org/10.1016/j.system.2008.02.004
McKay, T., & Plonsky, L. (in press). Reliability analyses: Estimating error in L2 research. In P. Winke & T. Brunfaut (Eds.), The Routledge handbook of second language acquisition and language testing. Routledge.
Moher, D., Stewart, L., & Shekelle, P. (2015). All in the family: Systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. *Systematic Reviews, 4*(1), 1–2.
Mozaffari, S. H. (2017). Comparing student-selected and teacher-assigned pairs on collaborative writing. *Language Teaching Research, 21*(4), 496–516. https://doi.org/10.1177/1362168816641703

Ortega, L. (2005). For what and for whom is our research? The ethical as transformative lens in instructed SLA. *Modern Language Journal, 89*(3), 427–443. https://doi.org/10.1111/j.1540-4781.2005.00315.x
Oskoz, A., & Elola, I. (2020). *Digital L2 writing literacies*. London, UK: Equinox Press.
Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes, 14*, 48–59. https://doi.org/10.1016/j.jeap.2013.12.001
Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition, 35*, 655–687. https://doi.org/10.1017/S0272263113000399
Plonsky, L. (2014). Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. *Modern Language Journal, 98*, 450–470. https://doi.org/10.1111/j.1540-4781.2014.12058.x
Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *Modern Language Journal, 100*, 538–553. https://doi.org/10.1111/modl.12335
Plonsky, L., & Gass, S. (2011). Quantitative research methods, study quality, and outcomes: The case of interaction research. *Language Learning, 61*, 325–366. https://doi.org/10.1111/j.1467-9922.2011.00640.x
Plonsky, L., & Kim, Y. (2016). Task-based learner production: A substantive and methodological review. *Annual Review of Applied Linguistics, 36*, 73–97. https://doi.org/10.1017/S0267190516000015
Ponterotto, J. G. (2006). Brief note on the origins, evolution, and meaning of the qualitative research concept thick description. *The Qualitative Report, 11*(3), 538–549.
Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching, 42*(2), 147. https://doi.org/10.1017/S0261444808005612
Rogers, E. M. (1998). *An investigation of the effects of collaborative, computer-mediated communication and non-collaborative, computer-assisted writing skills practice on L2 writing. [Unpublished doctoral dissertation]*. University of Arizona.
Rosenthal, R. (1979). The file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641. https://doi.org/10.1037/0033-2909.86.3.638
Selcuk, H., Jones, J., & Vonkova, H. (2019). The emergence and influence of group leaders in web-based collaborative writing: Self-reported accounts of EFL learners. *Computer Assisted Language Learning*, 1–21. https://doi.org/10.1080/09588221.2019.1650781
Sharifi, M., Rostami AbuSaeedi, A., Jafarigohar, M., & Zandi, B. (2018). Retrospect and prospect of computer assisted English language learning: A meta-analysis of the empirical literature. *Computer Assisted Language Learning, 31*(4), 413–436. https://doi.org/10.1080/09588221.2017.1412325
Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20*(4), 286–305. https://doi.org/10.1016/j.jslw.2011.05.010
Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: John Wiley & Sons.
Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition, 26*(3), 365–398.
Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*(1), 119–158. https://doi.org/10.1111/1467-9922.00179
Storch, N. (2013). Collaborative writing in L2 classrooms. *Multilingual matters*. https://doi.org/10.21832/9781847699954
Storch, N. (2019). Collaborative writing. *Language Teaching, 52*(1), 40–59. https://doi.org/10.1017/S0261444818000320
Strobl, C. (2014). Affordances of Web 2.0 technologies for collaborative advanced writing in a foreign language. *CALICO Journal, 31*(1), 1–18. https://doi.org/10.11139/cj.31.1.1-18
Su, F., & Zou, D. (2020). Technology-enhanced collaborative language learning: Theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, 1–35. https://doi.org/10.1080/09588221.2020.1831545
Suzuki, W., & Storch, N. (2020). *Languaging in language learning and teaching: A collection of empirical studies*. John Benjamins. https://doi.org/10.1075/lllt.55
Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal, 82*, 320–337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x
Swain, M., & Watanabe, Y. (2013). Languaging: Collaborative dialogue as a source of learning. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley Blackwell.
Torres, J., & Cung, B. (2019). A comparison of advanced heritage language learners' peer interaction across modes and pair types. *Modern Language Journal, 103*(4), 815–830. https://doi.org/10.1111/modl.12594
Van Lier, L. (2004). The semiotics and ecology of language learning. *Utbildning & Demokrati, 13*(3), 79–103. https://doi.org/10.1007/1-4020-7912-5
Visonà, M. W., & Plonsky, L. (2020). Arabic as a heritage language: A scoping review. *International Journal of Bilingualism, 24*, 559–615. https://doi.org/10.1177/1367006919849110
Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, MA: Harvard University Press.
Wang, Y. C. (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. https://doi.org/10.1080/09588221.2014.881386
Wigglesworth, G., & Storch, N. (2012). What role for collaboration in writing and writing feedback? *Journal of Second Language Writing, 21*(4), 364–374. https://doi.org/10.1016/j.jslw.2012.09.005
Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing, 21*(4), 321–331. https://doi.org/10.1016/j.jslw.2012.09.007
Wu, H. J. (2015). *The effects of blog-supported collaborative writing on writing performance, writing anxiety and perceptions of EFL college students in Taiwan. [Unpublished doctoral dissertation]*. University of South Florida.
Yeh, S. W., Lo, J. J., & Huang, J. J. (2011). Scaffolding collaborative technical writing with procedural facilitation and synchronous discussion. *International Journal of Computer-supported Collaborative Learning, 6*(3), 397–419. https://doi.org/10.1007/s11412-011-9117-9
Yim, S. (2017). *Digital literacy in academic settings: Synchronous collaborative writing among linguistically diverse students.[Unpublished doctoral dissertation]*. Irvine: University of California.
Zhang, M. (2021). Understanding L1 and L2 interaction in collaborative writing: A lexico-grammatical analysis. *Language Teaching Research, 25*(3), 338–359.
Zhang, M., & Chen, W. (forthcoming). Exploring the effects of assessment approaches in synchronous computer-mediated collaborative writing. Journal of Second Language Writing.
Zhang, M., & Crawford, W. J. (2021). Attention to form in collaborative writing: language-related episodes in L1 and L2 use conditions. *Language Awareness*, 1–19. https://doi.org/10.1080/09658416.2021.1998087
Zhang, M., & Plonsky, L. (2020). Collaborative writing in face-to-face settings: A substantive and methodological review. *Journal of Second Language Writing, 49*, Article 100753. https://doi.org/10.1016/j.jslw.2020.100753
Zhang, M., Akoto, M., & Li, M. (2021). Digital multimodal composing in post-secondary L2 settings: A review of the empirical landscape. *Computer Assisted Language Learning*, 1–28. https://doi.org/10.1080/09588221.2021.1942068
Zhu, W., & Mitchell, D. A. (2012). Participation in peer response as activity: An examination of peer response stances from an activity theory perspective. *TESOL Quarterly, 46*(2), 362–386. https://doi.org/10.1002/tesq.22
Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition, 38*, 553–586. https://doi.org/10.1017/S027226311500025X

Meixiu Zhang (Ph.D., Northern Arizona University) is Assistant Professor of Applied Linguistics at Texas Tech University, where she teaches courses in applied linguistics to both graduate and undergraduate students. Her primary research areas include L2 writing, corpus linguistics, and instructed second language acquisition. Her work can be found in *Journal of Second Language Writing, Language Teaching, Modern Language Journal, Language Teaching Research,Language Awareness, Computer Assisted Language Learning, Corpora, System, Register Studies*, and more.

John Gibbons, M.S., is a Ph.D. student in the Department of Literature and Languages at Texas A&M University-Commerce. He teaches Composition & Rhetoric and Linguistics courses, and his main research interests include second language writing, computer mediated collaborative writing, and composition and rhetoric. He was an assistant editor for Best of the Journals in Rhetoric and Composition 2019. He has taught English as a Foreign Language in Jordan, China, and Taiwan, as well as ESL in the US.

Mimi Li(Ph.D., University of South Florida), Ph.D., is Assistant Professor of Applied Linguistics/TESOL in the Department of Literature and Languages at Texas A&M University-Commerce. Her main research areas are second language writing and CALL. She has conducted research projects on online collaborative writing, computer-based instructor and peer feedback, gamification and vocabulary learning, and multimodal pedagogy in teacher education. Her work has appeared in *Journal of Second Language Writing*, *Computer Assisted Language Learning*, *Language Learning & Technology*, *Language Teaching*, *System*, *Computers & Education*, and *Computers and Composition*, among others.