



Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing

Heike Neumann*, Kim McDonough

Department of Education, Concordia University, 1455 De Maisonneuve Blvd. West, LB 579, Montreal, Quebec H3G 1M8, Canada

Abstract

In second language (L2) writing classrooms, prewriting discussions are one of the most commonly used collaborative activities (Fernández Dobao, 2012; Storch, 2005), yet there has been little research about their relationship to students' written texts. Recent L2 writing research has examined the textual features of co-constructed texts (e.g., Elola & Oskoz, 2010; Kuiken & Vedder, 2002; Storch & Wigglesworth, 2007), whereas the pretask planning literature has focused mainly on the effect of individual planning (e.g., Ellis & Yuan, 2004; Kroll, 1990; Ojima, 2006). The current study investigates the relationship between interaction during collaborative prewriting tasks and students' written texts in an English for academic purposes (EAP) course. The findings indicate that structured collaborative prewriting tasks elicited student talk about content and organization, but there is only a moderate relationships between these prewriting discussions and the students' written texts. Implications for the use of collaborative prewriting discussions in EAP contexts are discussed.

© 2014 Elsevier Inc. All rights reserved.

Keywords: Second language writing; Second language interaction; Collaboration; Group work; Prewriting tasks

With the advent of communicative language teaching, collaborative learning has assumed an important role in the second language (L2) classroom, with collaborative activities seen as one of the key characteristics of communicative L2 teaching. The inclusion of collaborative activities in the L2 classroom has been justified through reference to sociocultural theory, in particular Vygotsky's (1978) pioneering claims that speech is an essential part of human cognitive development, and that language and cognitive skills develop through interaction with others and the physical world. Interaction plays an essential role in knowledge-building by creating opportunities for learners to elicit help from experts or simply articulate steps in the problem-solving process through internal or external speech. While it was originally thought that these forms of scaffolding could only be provided by an expert through, for example, guided participation (Rogoff, 1990), numerous studies have shown that learners scaffold each other during collaborative activities in a wide variety of L2 contexts (Abadikhah & Mosleh, 2011; Alegría de la Colina & García Mayo, 2007; Donato, 1988, 1994a, 1994b; García Mayo, 2002; Ismail & Samad, 2010; Kim, 2008, 2009; Leaser, 2004; Suzuki & Itagaki, 2007; Swain & Lapkin, 2001; Watanabe & Swain, 2007, 2008).

* Corresponding author. Tel.: +1 514 848 2424x2443; fax: +1 514 848 4520.

E-mail addresses: heike.neumann@concordia.ca, hneumann@education.concordia.ca (H. Neumann), kim.mcdonough@concordia.ca (K. McDonough).

L2 writing researchers have focused more narrowly on the types of scaffolding that occur when learners collaborate to co-construct written texts (DiCamilla & Anton, 1997; Fernández Dobao, 2012; Gutiérrez, 2008; Kuiken & Vedder, 2002, 2005). For example, in a series of studies that investigated the use of collaborative writing in English as a second language (ESL) classrooms, Storch and Wigglesworth reported that peers routinely scaffold each other, especially when collaborative dynamics allow for the transfer of knowledge among group members (Storch, 2002a, 2002b, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009, 2012). In order to better understand the nature of collective scaffolding, learner interaction during collaborative writing activities in both face-to-face and online environments has been studied extensively. The analysis has largely focused on the occurrence of language-related episodes (LREs), which have been defined as “any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others” (Swain & Lapkin, 1998, p. 328). Fewer collaborative writing studies have documented how learners discuss other elements of written texts, such as their content or organization (Elola & Oskoz, 2010; Storch, 2005).

Besides examining the nature of collaborative dialogue, researchers have also analyzed co-constructed written texts, at varying stages of the collaborative writing process, in order to determine the benefits of collaborative as opposed to individual writing. These studies have found that collaboration has a positive impact on linguistic accuracy (Fernández Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). Storch (2005) found that collaboratively produced texts are not only linguistically more accurate and complex but also better in terms of content: When students wrote a text in pairs, the theses of their texts were more appropriate, and their texts included fewer unnecessary details. Elola and Oskoz (2010) reported that pairs planned more carefully before beginning to write, whereas individuals worked on the structure of the texts throughout the writing process. The authors, however, did not examine whether the different processes had an impact on the quality of texts in terms of organization.

Thus, research on collaborative writing has shown that learners discuss language, content, and organization, and their interaction is positively associated with text quality. This leads to the question as to why these types of activities are not commonly used in L2 writing classrooms. There are at least two practical concerns that instructors, especially at the tertiary level, may have in relation to these activities: group assessment and time constraints. Group assessment can take multiple forms (Roberts, 2006), but at the university level most commonly groups are assessed as a whole (Wigglesworth & Storch, 2009). For this reason, Kagan (1995) clearly finds group marks inappropriate because they inadequately capture the individual group members’ capabilities and raise issues in terms of fairness, reliability, and meaningfulness of these types of assessments. Strauss and U (2007) also discovered various challenges faced by instructors in the implementation of group assessments, which may have a negative impact on how fair group assessment can be. Simply put, teachers have difficulty determining what each individual learner contributed to a collaborative text. The second constraint relates to the time required for learners to complete collaborative writing assignments during class. As Storch (2005) reported, when completing the same writing task, pairs spent more time producing a text collaboratively than individuals working alone. In a university writing class, time is often at a premium, and instructors may feel they simply do not have the extra time that is necessary for students to write collaboratively, especially when producing longer texts.

In order to overcome these constraints, researchers have explored whether using collaborative prewriting discussions achieves similar benefits as collaborative writing. Along with peer review, prewriting discussions are one of the most common collaborative activities in the L2 writing classroom (Fernández Dobao, 2012; Storch, 2005). Less is known, however, about the nature of learner interaction during prewriting discussions or its relationship to text quality. Pretask planning research has largely examined individual tasks, rather than collaborative activities, and has investigated the impact of planning on oral performance (e.g., Ellis, 2009; Foster & Skehan, 1996; Ortega, 1999; Skehan & Foster, 2001; Wigglesworth, 1997). Few studies have investigated planning in L2 writing, and most explored the effect of individual planning, rather than collaborative planning, on writing (Ellis & Yuan, 2004; Kroll, 1990; Manchón & Roca de Larios, 2007; Ojima, 2006).

Some insight into the potential benefits of collaborative prewriting discussions is provided by Higgins, Flower, and Petraglia (1992), who examined whether group planning sessions help students in first language (L1) writing classes evaluate their ideas and create writing plans. They found that collaboration allowed students to reflect on their ideas, and this reflection positively contributed to the quality of their writing plans. However, the occurrence of reflection depended on the roles that the listeners assumed when the writers were explaining their plans (supportive listener or critic). Only if listeners challenged the writers to explain their ideas in more depth or evaluate their ideas did both partners—listener and writer—engage in reflection. Their finding parallels those of the LRE studies, which found that

pair dynamics influence the amount and type of learner talk that occurs during collaborative task performance (e.g., Kim & McDonough, 2011; Storch & Aldosari, 2013; Storch, 2002a, 2002b; Watanabe & Swain, 2007). However, Higgins et al. (1992) did not analyze the students' texts, instead focusing on the quality of learner scaffolding and the individual writing plans.

Besides examining learner scaffolding during collaborative prewriting activities, some studies have investigated their impact on text quality in both L1 (Sweigart, 1991) and L2 writing classrooms (Shi, 1998). Sweigart (1991) compared the effectiveness of teacher lecture, teacher-led discussion, and student-led discussion for promoting reading comprehension and improving the quality of students' writing. He found that students both gained a better understanding of the texts that they had to read for the tasks and wrote better texts themselves after having participated in the student-led discussion compared to the other two conditions. In L2 writing, Shi (1998) compared three prewriting conditions—teacher-led, student-led, and no discussion—but did not find a statistically significant difference in their effect on the quality of students' writing.

These studies provide insight into practical considerations in the design and implementation of collaborative brainstorming activities. Shi (1998) found that unstructured group planning activities were disorganized because “ideas popped up erratically one after another . . . connected by relatively free association” (p. 336), which left students with the challenge of ordering and evaluating those ideas before writing. In contrast, the teacher-led discussion both generated ideas and provided guidance about how to organize those ideas into a writing plan. Higgins et al. (1992) reported that prompting learners to evaluate and challenge their partners to defend or justify ideas was more effective at promoting reflection about purpose and audience than prompting learners to request information. The most effective materials for promoting sustained reflection were those that encouraged learners to evaluate and justify choices as well as propose alternatives.

The small number of L2 prewriting discussion studies has left a number of unanswered questions. In the L2 writing planning studies, it is worth noting that text quality is often used as an indirect measure for the effectiveness of planning, but the relationship between the quality of planning and text quality has not been directly investigated. Other than Higgins et al. (1992), which was conducted in an L1 classroom, previous studies have not focused on the quality of learner interaction in terms of content or organization, which may directly affect the quality of content and organization in students' texts. Furthermore, the design of tasks that can successfully guide L2 students during collaborative prewriting discussions also needs to be investigated.

The present study aims to contribute to this line of research by exploring an optimal design of prewriting tasks. It addresses two research questions and explores the related pedagogical implications: (1) What do English for academic purposes (EAP) students discuss during collaborative prewriting activities? and (2) Is there a relationship between students' collaborative prewriting discussions and their written texts? These research questions were investigated in two studies carried out with different students enrolled in two classes of the same EAP writing course. Study 1 focuses on naturally-occurring classroom discourse recorded as part of the pilot for a larger and unrelated independent study about structural priming (McDonough, Neumann, & Trofimovich, 2014; Trofimovich, McDonough, & Neumann, 2013) in order to identify what aspects of writing students discuss while carrying out collaborative prewriting activities (research question 1). Based on the findings of Study 1, the instructions and materials used for collaborative prewriting were modified and implemented with different students the following year with the aim of focusing student attention and encouraging reflection on content and organization during the prewriting tasks. Study 2 describes what students discussed when they carried out the revised prewriting activities (research question 1), and also examines the relationship between the students' prewriting discussions and their written texts (research question 2).

Study 1

Method

Participants and instructional context

The participants were 19 English L2 students (8 men and 11 women) who were enrolled in undergraduate or graduate degree programs in business (10), engineering (4), social sciences and humanities (3), and computers (2) at Concordia University in Montreal, Canada. As is typical of the ESL program, half of the students were permanent residents of Canada, while the other half were international students who had recently arrived in Canada on student visas. As a result, their residence in Canada ranged from four months to three years. They represented a variety of L1

backgrounds including Chinese (7), Arabic (4), French (3), Farsi (1), Romanian (1), Russian (1), Spanish (1) and Urdu (1). They had already partially met the university's English proficiency requirement for admission, which was a TOEFL iBT score between 75 and 89 or equivalent, but based on their performance on the university placement test, they were required to take the first of two EAP writing courses. The class was taught by the first researcher, and it met for two 2³/₄-hour classes per week for 13 weeks. It focused on the development of general rather than discipline-specific academic language skills. For this purpose, students reviewed and learned a range of vocabulary and grammatical structures useful for reading and writing academic materials. All of the writing tasks were theme-based and covered different academic topics, such as education and socialization.

Data collection

As part of the pilot for a larger, unrelated study about structural priming (McDonough et al., 2014; Trofimovich et al., 2013), the first researcher's EAP writing class was video- and audio-recorded over a 13-week semester. The purpose of the pilot study was to identify the types of collaborative activities that were regularly used in the EAP class in order to guide the design of materials for use in subsequent series of empirical studies about structural priming. A research assistant attended each class to video-record the instructor and to place 8 to 10 digital audio-recorders throughout the classroom to capture student interaction for the entire class period. The audio-recorders were spaced throughout the classroom in order to capture the various types of interactions (whole-class, pair work, and small group activities) that occurred in a typical class. If students moved their seats to carry out collaborative tasks, the research assistant who was video-recording the class checked to make sure that at least one audio-recorder was close enough to capture the voices of students in each pair or group. Following her normal classroom routine, the instructor allowed the students to self-select groups, which ranged in size from two to four students. Because it was beyond the scope of the pilot for the unrelated study, the researchers did not collect detailed, personal information from each student or solicit their writing assignments for inclusion in the analysis. All of these writing assignments were graded by the instructor using an analytical rating scale similar to the one used in Study 2 (see Appendix A),¹ and each assignment contributed 2.5% towards the students' final grade. Each assignment was worth a small percentage of the students' final grade because they were required to produce numerous assignments throughout the semester, for a cumulative value of 20% of the total grade.

Once the semester had ended, the audio-recordings were transcribed by research assistants. Although all of the students in the class ($N = 24$) had to carry out the activities because they were a regular part of the curriculum, they could decline to allow their data to be used for research purposes. Therefore, the research assistants only transcribed the voices of students who consented to participate in research. They consulted the video-recordings to identify which students were recorded on the various digital audio-recorders throughout the classroom in order to exclude non-consenting students from the transcripts. The researchers reviewed the transcripts to identify which class periods had included collaborative prewriting discussions, which lasted from 15 to 20 minutes. Only transcripts that had collaborative prewriting discussions with groups in which all group members had consented to participate in research were included in the analysis. After the researchers identified five relevant class periods and group transcripts, research assistants verified the transcripts by rechecking the audio-recordings and consulting the video-recordings as needed to ensure accurate identification of students in the prewriting discussions. Table 1 provides descriptive information about the classes included in the dataset.

Data analysis

Transcripts of the collaborating prewriting discussions were analyzed to identify student talk about content and organization (detailed description and examples of coding categories below). Because our focus was on student collaboration, interaction between an individual student and the instructor or between an entire prewriting discussion group and the instructor were not included in the analysis. Episodes involving the instructor rarely occurred and only if a student asked a question or requested assistance. Content episodes were operationalized as interaction in which students generated and proposed ideas. A content episode was defined as having one main idea along with any reasons

¹ The overall structure of the rating scale and the descriptors were the same in both studies. Whereas the instructor used the rating scale in Study 1 to provide students with holistic grades as part of the normal instructional routine, research assistants used the rating scale in Study 2 to generate subscores for content, organization, language use, and form/mechanics that could be analyzed separately.

Table 1
Target class periods.

Topic	Instructions for prewriting discussion	Writing assignment
Overcoming obstacles	Describe the person you have chosen to write about & organize your ideas	Write a paragraph about a person from your life who experienced an obstacle and describe how he/she overcame it
Socialization	Discuss how family, schooling peers, & media have influenced your socialization & refer to the textbook readings if needed	Write a paragraph in which you explain how one factor (family, schooling peers, or media) has most influenced you
Collective living	Discuss the benefits and costs of social living for humans & organize ideas	Write a paragraph that explains whether the benefits of social living outweigh the costs for humans
Education	Discuss the pros and cons of home schooling & organize ideas	Write a paragraph that explains your opinion about home schooling
Family living	Discuss changes in family size, or the effect of family size on women or children	Write a paragraph that describes changes in family life (size, role of women, or role of children)

or supporting details. Transitions between content episodes typically occurred in the form of questions (*okay, what else? so what's the next one? any other ideas?*) or suggestions (*okay so we go to the next one*). Following Higgins et al. (1992), all content episodes were further classified as reflective (involving explicit evaluation, consideration of alternatives, or justification) or non-reflective (having none of the features of reflective episodes). To be considered reflective, positive evaluation comments had to be more substantial than simple responses (such as *yeah* or *good*), which could simply function to move the conversation forward. Evaluation also involved recognition of a need or gap with the ideas that had been generated. Reflective comments that were a consideration of alternative plans involved a critical choice or comparison among options and were often signalled by *or* and *instead*. However, simple lists of options that were not discussed or explained at length were not coded as reflective. Lastly, reflective comments in the form of justification involved explanation of ideas and reasoning and were marked linguistically by conjunctions such as *because*, *so*, and *since*.

An example of a reflective content episode with evaluation is shown in (1), in which the students were talking about what individuals they could write about for the *Overcoming obstacles* assignment. S1 has explained that she might write about a Chinese man she had heard about from a friend who began studying English at age 36, mastered it within two years, and then received eight university degrees within 1.5 years. However, S8 evaluates this choice by referring to the instructions which specified that the individual should be someone that the students know personally.

(1) Reflective content episode

S1: yeah so in one term he got his diplomas ... in only one and a half years

S8: I don't think that's our teacher's meaning. I don't think the teacher want us write about somebody who was on TV. Just we know of.

S1: familiar?

S8: yeah

S1: She want us to write her a real person

S8: Yeah I think but I can't find a person worth to write about. I think most of my friends and my relatives are normal

In contrast to the reflective episode, (2) illustrates a non-reflective content episode in which the students were discussing the pros and cons of homeschooling. One student (S11) stated one potential disadvantage of homeschooling, and her partners agreed. After this short exchange, they continued stating and acknowledging additional disadvantages but did not elaborate or evaluate them.

(2) Non-reflective content episode

S11: Okay can we turn to the second one?

S1: Yes

S10: Yes

S11: Financial resources

S1: Yes we should have necessary supplies

S10: Okay

Table 2
Overview of coding categories.

Episode type	Description
Reflective content	Student talk about content and ideas that contains explicit evaluation, consideration of alternatives, or justification
Non-reflective content	Student talk about content and ideas without any evaluation, alternatives, or justification
Reflective organization	Student talk about the grouping, ordering, or linking of ideas that contains explicit evaluation, consideration of alternatives, or justification
Non-reflective organization	Student talk about the grouping, ordering, or linking of ideas without any evaluation, alternatives, or justification

Episodes in which students shared personal experiences related to the course themes that did not contain any evaluation, alternative plans, justification, or an explicit link to the writing assignment were coded as non-reflective.

Next, following previous collaborative writing research (Elola & Oskoz, 2010; Storch, 2005), organization episodes were operationalized as interaction involving the grouping, ordering, or linking of ideas. These episodes included discussion about (a) where specific ideas should be placed in outlines, charts, or tables, (b) what order ideas should be presented in writing, and (c) how links between ideas or between reasons and examples could be made or improved. Because the data did not include any organization episodes, an example from Storch (2005) is provided in (3).

(3) Organization episode

M: First we describe

C: Describe before first after that we should compare

M: Describe after that yeah after that we talk about . . . we describe this

(Storch, 2005, p. 159, C&M, lines 42–46)

All of the transcripts were coded by the second researcher while a subset of the data (43%) was coded by the first researcher for interrater reliability, and Cohen's kappa was .89. Disagreements were resolved through discussion and included in the analysis. Table 2 summarizes all the coding categories used during the data analysis.

Results and discussion

The presentation and discussion of results focuses primarily on the findings that informed the design of the collaborative prewriting tasks implemented in Study 2. Table 3 lists the total number of episodes and the mean number per group with standard deviations that occurred during the collaborative prewriting discussions by topic. Content episodes clearly dominated the students' discussions as no organization episodes occurred in the entire data set. Non-reflective content episodes were more frequent ($M = 5.60$; $SD = 4.53$) than reflective episodes ($M = 1.40$; $SD = 1.18$).

One possible explanation why non-reflective content episodes dominated the students' prewriting discussions is that this type of interaction is the most natural way to talk about the assigned topics. Unless students are strongly encouraged to critically reflect on their peers' ideas and comments, they may just opt to chat about the assigned topic without engaging critically with their own and their peers' ideas. The instructions for the prewriting discussion in Study 1 did not specifically ask students to provide each other with feedback, to consider alternatives, or to justify their opinions. For three of the five target classes, students were told to organize their ideas, but they preferred to exchange ideas about content rather than create a writing plan or outlines for the ideas they were discussing.

A number of qualitative findings also shed light on the nature of students' prewriting discussions. In general, their non-reflective content episodes belonged to one of two categories. The first type of non-reflective episodes involved students nominating and acknowledging ideas without actually engaging in any discussion or evaluation of those ideas. This type of interaction was illustrated previously in (2). The second type of non-reflective episodes involved a student sharing personal anecdotes without the other students engaging or pushing that speaker to evaluate or justify how the experiences related to the topic or could be developed as part of the writing assignment. This type of interaction is illustrated in (4), which occurred when students were discussing how their family and peers had affected their socialization. When S18 shares personal experiences about her relationship with her mother, her classmates, S6

Table 3
Content and organization episodes by topic.

Topic	Reflective content			Non-reflective content			Organization Sum
	Sum	<i>M</i>	<i>SD</i>	Sum	<i>M</i>	<i>SD</i>	
Overcoming obstacles (3, 8) ^a	3	1.00	1.00	10	3.33	.58	0
Socialization (2, 7)	2	1.00	1.00	26	13.00	8.49	0
Collective living (3, 9)	8	2.67	1.15	12	4.00	2.65	0
Education (3, 8)	3	1.00	1.00	16	5.33	3.79	0
Family living (4, 11)	5	1.25	1.50	20	5.00	3.56	0
Total	21	1.40	1.18	84	5.60	4.53	0

^a Number of groups, number of students.

and S19, are sympathetic and supportive listeners, but they do not push the speaker to make connections between the memory and their writing task, to clarify the main point the anecdote illustrates, or to evaluate whether it would be an appropriate supporting detail to include in her paragraph.

(4) Non-reflective content: Supportive listening

S18: But for me okay I'll tell you. For me, I think I'm a very sensible person because I was influenced by my mom. My mom was not – not that she was a bad influence on me, but she was really hard on me. So I become kind of like allergic to my mom.

S6: (laughs)

S18: So now that I'm a mom I don't want to do what she did

S19: I know what you're feeling

S18: You know? I became kind of like, okay, I love the way she brought me, you know she was okay, she was always taking care of me physically, like she was always giving me clothes and you know taking care of my school and things like that, but she was never there for me you know?

S6: I know what you mean . . . she was doing the mother's duty—

In brief, like Higgins et al. (1992), we also found that there was little evaluation when the students took on the role of supportive listener as opposed to critic (see episode in Example 4). Like Shi (1998), we also found that students generated a lot of ideas but did not spend as much time evaluating them and never talked about ordering them. In other words, when left to their own devices (unlike in Shi's teacher-led planning), the students successfully generated ideas but were not as successful at evaluating those ideas or organizing them into a writing plan.

These findings point to a number of areas where the collaborative prewriting activities could benefit from revision to make them more effective for generating student reflection on content and organization. First, the activities need to provide a narrower focus so that students can concentrate on generating and evaluating content, and organizing that content into a prewriting plan. Second, the writing assignment to be completed individually after the discussion should be included in the instructions. By providing the writing topic, students may understand more explicitly that the purpose of the prewriting discussions is to prepare for the writing assignment. Because the prewriting discussions were intended to lead students from a broader discussion of the topics to a writing topic that was not revealed until the end of the discussion, the students often spent more time trying to predict the writing assignment than discussing ideas. Third, the prewriting discussion activities should place greater emphasis on reflection so that students are encouraged to evaluate each other's ideas and push each other to justify ideas and choices. Finally, the prewriting discussions need to help students transition from the discussion of content to the organization of ideas in preparation for writing.

Study 2

Based on the findings from Study 1, the researchers created a structured prewriting discussion task, whose implementation and effectiveness are tested in Study 2. Besides considering the effectiveness of the revised task in terms of the students' discussion of content and organization, Study 2 also explores the relationship between the students' discussions and written texts. Thus, Study 2 addresses both research questions: (1) Do collaborative

prewriting discussion activities elicit talk about content and organization? and (2) Is there a relationship between students' collaborative prewriting discussions and their written texts? Although Study 2 was conducted with a different group of participants, the two groups of students were very similar: They had similar proficiency levels because they had met the university English language proficiency requirement, were placed in the same course based on an in-house placement test, and had similar L1 backgrounds. Besides the obvious reason that the participants in Study 1 could not participate in Study 2 because they had already completed the EAP class, asking the same participants to repeat the prewriting discussions and write paragraphs about the same topics would introduce practice as potential explanation for any differences in performance, as opposed to the new task materials.

Method

Participants

The participants in Study 2 were English L2 students ($N = 23$) who were enrolled in undergraduate degree programs in business (10), social sciences and humanities (7), engineering (4), and computers (2) at Concordia University in Montreal, Canada. They ranged in age from 18 to 46 years, with a mean age of 22.1 years ($SD = 5.8$), and reported speaking a variety of first languages including Chinese (12), Arabic (3), French (3), Farsi (1), Korean (1), Portuguese (1), Spanish (1), and Turkish (1). At the beginning of the semester, they reported having lived in Canada for less than one month up to four years, which corresponds to typical length of residence reported by international students in Canada on student visas and by students with permanent resident status in Canada, respectively. They had already partially met the university's English proficiency requirement for admission by submitting a TOEFL iBT score between 75 and 89 or equivalent and were required to take the same academic EAP writing class described in Study 1 based on their performance on the university placement test. The EAP class was taught by the first researcher, and none of the students in Study 1 participated in Study 2. The studies were separated by one year, during which time no changes to the class syllabus, materials, goals, or assessment procedures were made.

Materials

Structured prewriting tasks were designed for six writing assignments, five of which had been the focal point of the targeted class periods in Study 1. Informed by the findings of Study 1, all six tasks contained the same three parts: (a) the writing topic, (b) a section about generating and evaluating ideas, and (c) a section about selecting and organizing ideas into a writing plan. Both the content and organization sections included individual brainstorming followed by a collaborative discussion in which students shared their ideas and presented their writing plan. The instructions emphasized that students should provide feedback about the quality of the ideas and the writing plan to their group members. [Table 4](#) presents an overview of the topics for the six writing assignments and the general instructions provided before the students began each structured prewriting task. The first structured prewriting task is provided in [Appendix B](#) as an example.

Procedure

For all six topics, the same procedure was followed in class. The structured prewriting tasks were distributed to students, and the teacher explained the instructions. As in Study 1, the instructor followed her normal classroom routine by allowing the students to self-select groups, which again ranged in size from two to four students. Although students tended to carry out the prewriting tasks with the same classmates for all six tasks over the semester, there was some variation in group composition due to student absences. Each group worked independently through the individual and collaborative components of the content and organization sections, which took about 20–30 minutes of class time. Although the structured prewriting tasks took more class time to complete than the prewriting discussions used in Study 1, the amount of discussion time (15–20 minutes) was the same. The additional time required to implement the structured prewriting tasks was due to the individual generation of ideas in Part 1 and the individual planning in Part 2. As each group finished, students left class and wrote the assignments as homework (topics 1–2) or began writing their assignments to be submitted before the end of the class period (topics 3–6). Students had one hour to complete the assignments written in class, but few students took the full hour to finish writing. It was not possible to determine how much time students spent writing the two texts they produced at home. Assignments completed as homework were submitted electronically, while in-class writing assignments could be completed electronically or by hand.

Table 4

Prewriting tasks: writing topics and instructions.

Topic	Writing assignment	Instructions
Overcoming obstacles	Write a paragraph about person from your life who experienced an obstacle, and describe how he/she overcame it	Review the list of obstacles collected on the board, and think of people in your life who faced and overcame similar obstacles
Socialization	Write a paragraph in which you explain how one factor (family, schooling, peers, or media) has most influenced you	Review the text on these factors, consider how each factor influenced you, and think of concrete illustrative examples
Collective living	Write a paragraph that explains whether the benefits of social living outweigh the costs for humans	Review the costs and benefits for animals discussed in the textbook, and reflect whether they apply to humans
Education	Write a paragraph that explains your opinion about home schooling	Review texts on homeschooling, and list reasons for and against it
Family living	Write a paragraph that discusses changes in family size (nuclear/extended or number of children)	Take notes on how family size changed, and list reasons for and consequences of the change
Before/after vaccines	Write a paragraph that compares and contrasts human life before and after the invention of vaccines	Use a Venn diagram to brainstorm differences and similarities about the time before and after vaccines

Data analysis

Research assistants transcribed and verified the audio-recordings of the students' structured prewriting discussions, and the transcripts were analyzed to identify episodes in which students discussed content and organization (see Table 2 for an overview of the coding categories). For content episodes, the same two categories (reflective and non-reflective) used in Study 1 were adopted, which were previously illustrated in Examples (1) and (2). As in Study 1, organization episodes included discussion about (a) where specific ideas should be placed in outlines, charts, or tables, (b) what order ideas should be presented in writing, and (c) how links between ideas or links between reasons and examples could be made or improved. In anticipation that the structured prewriting task would elicit organization episodes, a distinction was made between reflective and non-reflective episodes. Similar to reflective content, reflective organization was defined as having explicit evaluation, consideration of alternatives, or justification, while non-reflective organization lacked any of these characteristics. In the reflective organization episode with evaluation in (5), S3 was describing her writing plan by listing her reasons and the examples she would use to support each example. When she finished, S4 suggested that a different example was needed. The focus of the episode was not the content of S3's idea (child-parent relationships or family size) but the evaluation of the link between the reason and the example.

(5) Reflective organization episode

S2: Start start you were the first

S3: Okay ... I just see if I have the right structure

S2: Okay

S3: ... The second reason is less freedom if we live in a group. The example I will put then is children and parent relationship or a people live in a big family or big society yeah

S2: Okay I think it's good

S3: Thank you

S4: I think you need a more specific example there

In contrast, in the non-reflective organization episode in (6), S17 states his writing plan, but the group members do not provide any evaluation or alternatives or request any justification for the plan.

(6) Non-reflective organization episode

S14: So what's your plan?

S17: Mine? I just give examples to benefits. That's it. Then, my opinion, two examples of benefits ...

S14: Yeah

S17: Right?

S14: Yeah

S12: Okay I finished it's time to give for a break.

Following a training session with the second researcher, a research assistant coded all of the prewriting discussion transcripts. The training consisted of a review of the coding categories, collaborative coding of one transcript, and independent coding of one transcript followed by a comparison and discussion. For interrater reliability, the second researcher independently coded a subset of the data (16%) and Cohen's kappa was .91. All disagreements were resolved through discussion and included in the analysis.

The students' written texts for all six assignments were rated by two trained raters using an analytical rating scale (see Appendix A) with subscores for content, organization, grammar/vocabulary, and mechanics/form that is based on the rating scale used in the ESL program (see also Footnote 1). The raters were trained by the first researcher using benchmark papers employed for the regular rater training session in the ESL program, and both raters rated all texts for the first assignment. The researcher and raters then met again to discuss and compare ratings for 50% of the texts for the first assignment. The two raters then independently rated the remaining texts, and interrater reliability, assessed using a two-way mixed average-measures intraclass correlation coefficient, was .74. The means of the content and organization subscores assigned by the two raters for each text were calculated, and these means were used for subsequent analyses. Spearman's rank correlation coefficients between the total number of reflective content and organization episodes each student participated in and the students' content and organization subscores as well as the overall score for each assignment were calculated. There was no statistically significant difference in the students' content, organization, or total scores by topic or by writing condition, i.e., in class or at home (Wilk's $\Lambda = .044$, $F(2,15) = 2.87$, $p = .29$).

Results

The first research question asked whether the structured prewriting tasks elicited student talk about content and organization. Table 5 provides the sum and mean number of reflective and non-reflective content and organization episodes with standard deviations that occurred during each structured prewriting task. For content episodes, non-reflective episodes occurred more frequently than reflective episodes across all six topics. There was more variation in the distribution of organization episodes, with two topics (*Collective living* and *Vaccines*) eliciting more reflective episodes than non-reflective episodes.

When compared to the Study 1 data, the mean number of reflective content episodes increased (1.40 and 2.45, respectively), as did the mean number of non-reflective content episodes (5.60 and 8.28, respectively). In addition, the increased occurrence of organization episodes, both reflective and non-reflective, is noteworthy considering that no organization episodes occurred in Study 1. In short, it appears that the structured prewriting discussion task was more effective at eliciting reflective content and organization episodes than the less structured activity described in Study 1. Independent-samples *t*-tests indicated that there were significantly more non-reflective content episodes in Study

Table 5
Content and organization episodes by topic.

Topic	Reflective content			Non-reflective content			Reflective organization			Non-reflective organization		
	Sum	<i>M</i>	<i>SD</i>	Sum	<i>M</i>	<i>SD</i>	Sum	<i>M</i>	<i>SD</i>	Sum	<i>M</i>	<i>SD</i>
Overcoming obstacles (7, 23) ^a	17	2.43	2.70	61	8.71	2.29	5	.71	.95	9	1.29	1.38
Socialization (6, 20)	17	2.83	2.79	54	9.00	5.14	8	1.33	2.16	12	2.00	2.19
Collective living (7, 23)	31	4.43	4.93	64	10.00	4.80	5	.71	1.11	3	.43	.79
Education (7, 22)	5	.71	.95	41	8.00	3.61	3	.43	1.13	4	.57	.79
Family living (7, 20)	17	2.43	2.88	47	6.71	3.25	1	.14	.38	1	.14	.38
Vaccines (6, 21)	11	1.83	.75	30	7.17	3.60	6	1.00	2.00	3	.50	.84
Total	98	2.45	2.94	331	8.28	3.78	28	.70	1.34	32	.80	1.27

^a Number of groups, number of students.

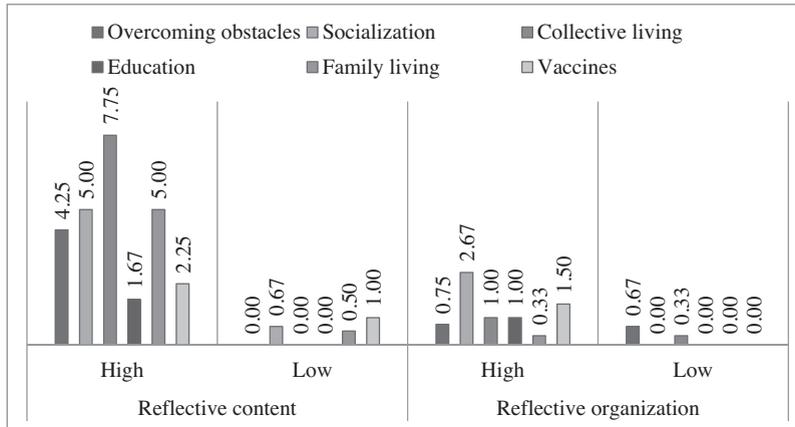


Fig. 1. Episodes by topic and engagement.

2 [$t(47) = 2.29, p = .027, d = .68$], and the increase in reflective content episodes approached significance [$t(47) = 1.95, p = .058$, equal variance not assumed, $d = .51$]. Because Study 1 did not include the *Vaccine* prewriting task, those data were excluded from Study 2. The organization episodes were not compared using inferential statistics because none had occurred in Study 1.

However, the larger dataset of Study 2 contained numerous groups that showed little engagement in the task. As a result, the aggregate group data shown in Table 5 does not capture the variability shown across groups. While students in some groups simply took turns stating their ideas and writing plans without engaging in any discussion or evaluation, students in other groups were engaged in the task of evaluating both ideas and organization and providing each other with feedback. In order to gain a more nuanced view of the structured prewriting tasks discussions, the groups were classified as high and low reflectors, based on the median occurrence of reflective content episodes per group (2.00). The discussion groups that engaged in reflective episodes at rates at or above the median were classified as high reflectors, and the groups that had only one or no reflective episodes were classified as low reflectors. As shown in Fig. 1, in contrast to the low reflector groups, the high reflector groups critically engaged during the discussions and accounted for almost all of the reflective content and organization episodes. The divergent performance between high and low reflectors highlights the variability inherent in L2 classrooms as well as the challenges instructors face in communicating the purpose of pedagogical activities so that students perceive the value of engaging in those activities.

The second research question asked whether there is a relationship between the collaborative prewriting discussions and the students' written texts. In order to identify this relationship, Spearman's rank correlations coefficients were calculated between the total number of reflective content episodes for each student and the content subscore and the overall score for each assignment. Since the number of reflective organization episodes was low for the data set, correlation coefficients for organization were not calculated. Table 6 presents the means and standard deviations for the content subscores and overall scores by topic and the Spearman's ρ correlations coefficients between the total number of reflective content episodes and (a) the content subscores and (b) overall scores for each topic. For the content subscores, the moderate correlations with the total number of reflective content episodes for the *Socialization* and *Collective living* topics were statistically significant. For the overall paragraph scores, the moderate correlations with the total number of reflective content episodes for the *Socialization*, *Collective living*, *Education*, and *Family living* topics were statistically significant.

To better understand the connection between the occurrence of reflective content episodes during the prewriting discussions and the students' written texts, excerpts from the transcripts and texts were examined. The data from one group of four students during the *Collective Living* unit illustrates the variability in these moderate relationships. The students in (7) worked together throughout the entire semester and regularly generated reflective content episodes, but they did not consistently incorporate the feedback they received during the discussions into their written texts. For example, S13 received evaluative feedback that had potential to improve the content of her text in (7).

Table 6

Means, standard deviations, and Spearman's ρ with reflective content episodes for content subscores and overall scores by topic.

Topic	Content subscore			Overall score		
	<i>M</i>	<i>SD</i>	Correlation	<i>M</i>	<i>SD</i>	Correlation
Overcoming obstacles	7.75	.85	.09	26.9	2.71	.29
Socialization	8.06	.95	.50*	28.6	3.00	.47*
Collective living	7.68	.87	.47*	27.32	2.73	.44*
Education	8.51	.59	.27	29.80	1.79	.53*
Family living	7.88	.89	.40	27.39	2.81	.45*
Vaccines	7.64	1.03	.23	27.14	2.65	.15

* $p < .05$.

(7) Reflective content episode without incorporation

S13: And, last point I found is the—costs is food. If the population goes up the amount of food gets less and less for each individual, so it is both in humans and animals.

S20: I the third one is, um—can be more specific.

S9: Yeah.

S13: Ok.

However, she did not follow the advice to make this point more specific. Instead, she simply stated it without any support, whereas an explanation for the reasons of this food shortage or its consequences on humans was called for. Another member of this group, S18, similarly received feedback during the prewriting discussion, as shown in (8), when he solicited help after having presented his content ideas on both the benefits and costs of social living.

(8) Reflective content episode with incorporation

S18: But which one can you work? Benefits or costs?

S9: I think you have really good examples of costs.

S18: Yeah?

S20: A lot of.

S13: I think you should write about costs.

In his text, S18 focused on the two costs of social living he had explained in detail, following the group's recommendation. In his text, however, S18 did not include all the detailed information which he had expressed to his group but left part of it out. The idea of competition for food and jobs was well explained in his text, but the idea how social living might be exploited by some people was not explained in as much detail in writing as it had been during the discussion.

Additional insight into the variability of the relationship between the prewriting discussions and written texts is provided by student groups that rarely engaged in reflective episodes. For low reflector groups, the most common interactional pattern was for a group member to deliver a monologue about their ideas followed by "that's all," which served as a cue for the next person to start another monologue. The discussions concluded when all had finished describing their ideas, and no group members offered or solicited feedback. In their written texts, students then reproduced what they had presented to the group. For example, S4 explained three benefits of social living for animals (cooperation, defense of territory, and raising young) during the discussion and also wrote about these despite the inappropriateness of the content (the topic focused on humans not animals). Likewise, S22 focused on the two benefits for humans (shared resources and cooperation) that he had explained to the group. Unlike in S4's case, his ideas were appropriate for the assigned topic.

Although it was not the main focus of the study, the students' perceptions about the structured prewriting discussions were elicited through a questionnaire administered at the end of the semester. The questionnaire contained five Likert-scale items (1 = not helpful; 9 = very helpful) that asked students to indicate how helpful it was to collaborate with their peers for different components of the writing process (understanding task requirements, brainstorming main ideas, generating supporting details, evaluating ideas, and organizing ideas). The questionnaire

also included three, open-ended questions that asked students to elaborate on what aspects of structured prewriting discussions were helpful/unhelpful, what English skills they developed through the discussions, and whether they preferred individual or collaborative prewriting activities.

Overall, the students reported positive perceptions of the structural prewriting discussions, with their highest rankings of 7.24 for brainstorming and supporting details ($SD = 2.07$ and 1.55 , respectively). They believed the tasks were helpful for understanding the assignment, audience, and purpose (7.05 , $SD = 1.56$) and for evaluating which ideas were appropriate for the assignment requirements (7.00 , $SD = 1.58$). Their lowest rating was 6.57 ($SD = 2.20$), which was for organizing ideas into a prewriting plan. When asked if they prefer to work alone or with classmates to carry out prewriting activities, 16 of the 22 students stated that they preferred to work with classmates because they can generate more ideas, exchange different viewpoints, clarify information they had not understood, and get suggestions from their peers. For the six students who preferred to work alone, they mentioned that they had greater concentration and less stress or time pressure when working individually. Two students believed that working alone was preferable because their classmates viewed the discussion tasks as a way to appropriate other student's ideas.

When elaborating on the aspects of the prewriting discussions that were helpful, 21 of 22 students mentioned the general topics of ideas and content most often, specifically gathering more ideas to write about, selecting examples, developing a better understanding of the topic, and being exposed to critical views that lead them to find more relevant examples. Only eight students mentioned organization, with five of them stating that the prewriting discussions helped them improve the organization of their ideas. However, three students stated that the organization sections of the prewriting discussions were not helpful, and that this component of prewriting could be best done individually. In sum, the questionnaire data largely confirm the results, as students talked about content most frequently during the prewriting discussions and also reported that the prewriting discussions were most helpful for developing their ideas. Similarly, the students talked about the organization of the ideas less frequently during the prewriting discussions, and fewer students reported that prewriting discussions were helpful for organizing their ideas.

Discussion

In terms of the first research question, which asked about student interaction during prewriting discussions, the structured prewriting tasks used in Study 2 were more effective in engaging L2 writers in critical evaluation of their ideas and organization than the naturally-occurring discussions found in Study 1. For all but one topic, the prewriting discussion tasks led to a doubling of reflective content episodes compared to Study 1. By modifying the instructions and the format of the task materials, the purpose of giving feedback and evaluating group members' ideas and writing plans was more explicit. The revised task materials encouraged discussion of organization, which led students to discuss how to structure their written texts in Study 2. Including the writing topic in the task materials also reduced the amount of time students spent speculating about the writing topic and encouraged a greater focus on generating and evaluating ideas and organizing those ideas into a writing plan.

One possible explanation for this finding is that the structure of the task led students towards reflection on content and the organization of their ideas. As seen in Study 1, a natural tendency for students was to avoid the socially and cognitively more challenging task of evaluating their peers' ideas. Socially, this task is challenging because they may have to critique or question their peers' comments; cognitively, the task is challenging because they cannot simply listen but have to critically engage with their peers' ideas in order to evaluate their suitability for the assigned topic. The structured prewriting task may have supported students in two ways. First, the generation and the evaluation of ideas were separated into two distinct steps (Parts 1a and 1b of the task), which possibly allowed students to be more critical in their evaluation than they would be during a brainstorming activity. Secondly, the task and the instructions push students towards reflection and evaluation in a number of places (e.g., the written instructions and the feedback component of the diagram in Part 1), which may be necessary as reflection did not occur frequently in the naturally occurring, unstructured discussions of Study 1.

Similar to [Storch's research \(2002a, 2002b\)](#), this study also found variation in group dynamics, which appeared to impact the occurrence of reflective content episodes. Some groups adopted what [Storch \(2002a\)](#) called a collaborative approach, in which students volunteer feedback and actively seek answers to their questions. Examples (7) and (8) illustrated collaboration in the form of giving and soliciting feedback, and (9) further shows how groups with collaborative dynamics worked together to develop ideas. After S18 shared his ideas about the benefits of social living,

the group members contributed additional examples that he could use to support the main idea that the advantages of collective living outweigh the disadvantages.

(9) Collaborative group dynamics

S13: Uh in natural disasters—because you are living in a community, community forms little groups and helps people in earthquakes, for example.

S18: Yeah.

S13: You can use maybe, or flood.

S20: Yeah, is good example.

S13: Or for example, there is Red Cross and it, it, um . . .

S9: Yeah, UNICEF?

S13: Yeah, UNICEF.

S9: Yeah, UNICEF, Red Cross. . . . [S13 explains UNICEF]

S18: Ok.

The finding that a collaborative approach facilitated greater engagement with the ideas of others through evaluation and reflection parallels the findings of collaborative dialogue studies that focused more narrowly on LREs and also reported greater engagement with LREs among pairs and groups with a collaborative approach (e.g., Kim & McDonough, 2008; Storch, 2007, 2008, 2013; Storch & Aldosari, 2013; Watanabe & Swain, 2007).

When students collaborate, they can improve the quality of their ideas. For example in (10), S1 explained his ideas for the *Collective living* writing assignment during the structured prewriting discussion and was challenged by his peers to modify his ideas since he focused on the benefits of collective living for animals—as discussed in the article in the course textbook—as opposed to the benefits for humans, which was the assigned writing topic.

(10) Challenging the suitability of ideas for assignment

S1: Uh, the benefit—the first benefit it’s about capture food. It will be more easy to capture food, um, the example is the . . . black-headed grouse and lions. Yeah, it’s mentioned in the articles.

S16: Ok, develop for the human. Because in your paragraph—

S14: Yeah your paragraph you just—

S16: —you will talk only about this.

S14: Focus on human.

S16: This . . .

S1: Oh, example of human?

Later on during the same discussion, S1 reframed his ideas to focus on the benefits for humans and used these ideas in his assignment. In contrast, another group’s discussion about the same topic did not elicit the quality feedback illustrated in (10). Those students discussed the costs and benefits of collective living for animals and included details in their paragraphs about animals even though the writing assignment was to focus on the costs and benefits of collective living for humans.

The quality of feedback with which students provided each other also has an impact on the ideas they selected for the writing assignments. In one group that consistently had high numbers of reflective episodes, students provided each other with thoughtful feedback during the structured prewriting discussion for the *Socialization* topic. All three students shared their ideas on how the four factors that can influence a child’s socialization (family, schooling, peer group, and mass media) had influenced them personally. Following each student’s presentation of ideas, the peers provided insightful comments on which factor would be the best to choose based on the examples and information provided by each student. Excerpt (11) from their discussion illustrates their interaction.

(11) Evaluating ideas and examples

S18: Actually you have a lot of examples but

S13: A pile of examples

S20: Okay. So—

S18: But I prefer the schooling, because it's maybe like the . . . the one which give you the most. Because you learn uh . . . with friends, with . . . uh yourself about . . . knowledge.

S20: Oh yes.

S18: So I think it's better.

S13: Mm hmmm. It has different aspects like, different situations and different people, it is it is the . . .

S18: The more complete, I think.

S13: Mm? Yeah! It's a whole.

S20: Yeah, it's a whole thing.

S18: yah but, all examples are . . . very good.

S20: Yeah I can use, what my parents told me to practice in school.

S13: Mm hmmm.

S20: And . . . yeah. And . . . people will say that school is like a small society.

S18: Yeah.

S13: It *is*.

S18: It is a small society.

S20: Okay. I will take this one

S20 chose the idea recommended by her peers during the prewriting discussion for inclusion in the written assignment. This group's ability and willingness to provide each other with feedback contrasted with other groups, where students focused more on the exchange of ideas or information rather than the evaluative aspect emphasized in the task and the instructions. The interaction in excerpt (12) from another group's prewriting discussion on *Socialization* illustrates the lack of evaluation.

(12) Discussion without evaluation

S19: Alright my family my father is a businessman,

S2: Mm hmmm?

S19: So he – he gives me . . . the patience and serious and uh a model. Like . . . uh he always deal with his problem in the company so, I see it, I do it. So he teach me a lot.

S17: Yeah.

S19: I also got three sisters so . . . it's why I got patience.

S11: Three sister . . .

S17: Three sisters!

S19: Yeah. Three older sisters

S11: Is . . . blood relationship?

S17: Three older sisters

S19: Yeah. Three older sisters

S11: Three older sisters

S2: A large family

S17: So, sister or cousin?

S19: Okay, I got three older sisters, so they give me patience and a model to teach me, in the life, how to deal with uh . . .

S17: Problem?

S19: Yes

S2: Problems.

S19: So that's it.

S17: That's a . . . good example.

S11: Perfect!

Although his peers engaged S19 in a discussion about his family, they did not evaluate whether he had provided sufficient details to support his main idea. This contrasts with the feedback S20 received from her peers, which focused specifically on the quantity and quality of the details she provided to support her ideas.

In contrast with the collaborative approach reflected in (9), (10) and (11), other groups were less likely to engage with each other's ideas, as can be seen in (12). The group that generated episode (12) worked together consistently throughout the semester and generated only one reflective content episode across all six discussions. The students in this group appeared to consider the prewriting discussions as an opportunity to joke around instead of engaging with the task. In their discussion of *Overcoming obstacles*, a student suggested that another group member was a good example for this topic due to his language problems. Later on, another student proposed himself as a good example because he could not think of an actual person who overcame obstacles in their life. Even when a group member presented an inappropriate idea in the *Collective living* discussion (i.e., writing about animals when the topic was about humans), the other students simply responded with superficial comments such as "very good, perfect." Although constructive criticism or a challenge to tailor the topic to the writing assignment would have been more appropriate, these group members did not volunteer feedback or ask or answer questions. One possible explanation is that the students may have been overly challenged by the task to provide each other with useful feedback and therefore opted to avoid the task completely.

In terms of the second research question about the relationship between prewriting discussions and written texts, the findings were less straightforward. Higgins et al. (1992) found a link between the amount of reflection and the quality of writing plans. Based on their findings, we expected a positive correlation between the occurrence of reflective episodes during the prewriting discussions and the quality of written texts. However, only two writing topics revealed the expected relationship between reflective content episodes and content subscores. In other words, more reflection during the prewriting discussion did not necessarily lead to better content of written texts. This may in part be due to the fact that the relationship between students' writing plans and their ultimate texts may be fairly indirect, as Higgins et al. (1992) cautioned. Plans may change while students are writing their texts, especially if they are writing outside of class, as was the case for Topics 1 and 2. Furthermore, students may not necessarily integrate their peers' suggestions into their written texts, especially if they do not believe that their peers have sufficient knowledge and expertise to help them improve their writing. Some students may prefer to delay incorporating peer feedback until they have received instructor feedback that "validates" the peers' suggestions. The relationship between prewriting discussions and overall text quality was more consistent, with four out of the six topics having significant moderate correlations. A possible explanation is that students whose texts received higher scores may simply be better at engaging in reflective discussions and discussing their writing, i.e., good writers are good at talking about writing. Consequently, they are more engaged in discussions with their peers about writing although this speculation requires further empirical validation.

Although this study, like previous research (e.g., Elola & Oskoz, 2010; Fernández Dobao & Blum, 2013; Shehadeh, 2011; Storch, 2005) has found that students' perceptions about the usefulness of collaborative writing activities are generally positive, especially in terms of generating and evaluating ideas, researchers have also emphasized that not all learners can be expected to or forced to collaborate (Storch, 2002a; Watanabe & Swain, 2007). That is why some researchers allow students to choose to work collaboratively or individually (e.g., Storch, 2005; Storch & Wigglesworth, 2007). The current study, however, required all students to work in groups during the prewriting tasks, as other studies have done (e.g., Fernández Dobao, 2012, 2014; Fernández Dobao & Blum, 2013; Shehadeh, 2011; Wigglesworth & Storch, 2009). The varying degrees of engagement observed in this study may be at least partially explained by differences in students' beliefs about how beneficial these prewriting discussions are for their own writing. Students who believe that collaborative prewriting discussions are useful may be in a position to exploit the learning opportunities they provide and benefit from them. However, if students do not see how collaborative discussions can improve their individual writing skills, they may be less inclined to participate actively. It may be useful to provide students with a choice between collaborative and individual prewriting activities, which suggests a promising avenue for future research.

Besides students' willingness to participate, another factor that may contribute to the quality of prewriting discussions is the topic of the writing task. Across both studies, some topics generated more reflective content and organization episodes than others. However, as indicated above, there were no statistically significant differences in the students' scores across topics. For example, in both studies the topics of *Collective living* and *Family life* generated the most reflective content episodes, while *Vaccines* elicited the most organization episodes. *Collective living* may have been challenging for students because it required extrapolation from animal behaviour to human society. *Family life*, on the other hand, may have posed challenges because students had to describe a historical change, explain a reason for the change, and articulate a consequence. This required consolidation of three

distinct but related pieces of information into a relatively short text. Although the *Vaccine* topic provided a relatively straightforward prompt (to compare and contrast aspects of life before and after the invention of vaccines), the students had to evaluate which organizational pattern of those presented in class was most appropriate for their ideas.

Additional factors which may have impacted the findings are the amount of feedback that students need in order to improve their writing skills and the students' ability to provide useful feedback to their peers. Some students may have been better at providing their peers with feedback about ideas and organization. In some ways, this is related to the idea discussed above that good writers may be better at talking about writing. Some students may simply have a better ability to critique their peers' ideas and provide them with valuable feedback, which led to more reflective content episodes. Also, some students may require more detailed feedback or in a different format (i.e., clearly laid-out individual written feedback versus informal/spontaneous oral feedback in a group setting) in order to truly benefit from and use it in their own writing. As a result, students who prefer individual or written feedback may not have been able to improve their texts based on the spontaneous, oral feedback provided in reflective content episodes.

Contributions and suggestions for further research

This study makes two contributions to the previous collaborative prewriting research that have pedagogical implications for the L2 writing classroom. First, students in L2 academic writing classes can be successfully encouraged to evaluate ideas when provided with structured prewriting tasks similar to the ones used in this study. These structured tasks had a positive impact on the amount of reflection that occurred during group discussions, thereby increasing their potential impact on the quality of students' written texts. Because the structured prewriting tasks are flexible, they can easily be adapted to other kinds of written texts and genres although their usefulness for generating reflection and evaluation for these other tasks would have to be confirmed by future research. Secondly, as other research on collaborative tasks in the L2 classroom has found, these tasks work best with groups that adopt a collaborative approach. It is unclear, however, what makes some groups adopt a collaborative approach, whereas others do not. [Storch's \(2013\)](#) research on LREs during collaborative writing tasks has led to recommendations to model collaborative interaction, provide communication strategies for group interaction, and monitor group interaction, but further research is needed to confirm whether these suggestions also hold true for structured prewriting tasks. Future studies might also explore whether an individual student benefits more from the experience of providing feedback to peers, as compared to receiving feedback or observing two peers engage in reflective episodes.

Whereas structured prewriting tasks were successful at engaging students in reflective discussions about the content and organization of their texts, the relationship between those prewriting discussions and text quality was more nuanced. Other factors, such as students' perceptions about the value of collaboration, the quality of peer feedback, the writing topics and assignment guidelines, may impact this relationship. Our future research aims to explore the factors that mediate the relationship between collaborative prewriting discussions and individual text quality and address the limitations of the current study. By expanding beyond the current sample of relatively high proficiency students in an EAP context, future research should explore the contribution of peer collaboration to writing development in other contexts (such as EFL) and with writers from different proficiency levels (such as pre-academic writers). Due to the potential impact of additional factors, future research should examine the relationships between prewriting discussions and text quality in studies that control for writing topic and written text type. And finally, future studies should control for student preferences in prewriting activities (i.e., collaborative versus individual) in order to identify whether combining students' preferred approach to prewriting tasks influences the benefits gained from prewriting as well as text quality.

Acknowledgements

This research was supported by funding from the Social Studies and Humanities Research Council of Canada (SSHRC), the Canada Research Chairs program (CRC), and the Fonds québécois de la recherche sur la société et la culture (FQRSC). We would like to thank the research assistants for their help with data collection and transcribing: Zachery Alderton, Yuan Chen, Rebecca Dorner, Jennifer Foote, Liana Jalyayan, Paula Kielstra, Jihoon Kim, Maxime Lavalée, Pham Quynh Mai Nguyen, Xuan Qin, and Fernanda Soler.

Appendix A. Paragraph evaluation grid

	Content	Organization	Grammar and Vocabulary	Mechanics and Form
10 9	Above standard C Topic sentence clear & focused C Strong, convincing support C Specific secondary support C Appropriate conclusion	O Logical organization & effective sequencing O Effective & varied use of transitions	G Good variety of sentence types G No major problems with sentence combination G No errors in phrase structure/parallelism G Few errors (1–5) in grammatical forms V Good range and variety in vocabulary V Few errors (1–2) in word choice, word form, or idiomatic phrasing V Precise and effective word choice and register	M A few minor errors in spelling, punctuation, and/or capitalization (1–2) but meaning is clear F Proper paragraph form F Neat presentation
8 7	Standard C Topic sentence present; controlling idea imprecise C Mostly relevant and appropriate support C Occasional minor problems with depth of development and/or unity C Conclusion present	O Mostly logical organization & effective sequencing O Mostly effective & varied use of transitions	G Adequate variety of sentence types G One or two sentence combination problems G Few errors in phrase structure/parallelism G Occasional errors (6–12) in grammatical forms V Adequate range and variety in vocabulary V Occasional errors (3–6) in word form, word choice, or idiomatic phrasing V Meaning rarely unclear	M Several errors in spelling, punctuation, and/or capitalization (3–6) but meaning is mostly clear F No indentation F Mostly neat presentation
6 5	Approaching standard C Controlling idea not evident in topic sentence OR topic sentence not present/appropriate C Some supporting points vague, insufficient, unconvincing, and/or off-topic C Problems with unity and/or focus or development of topic C Conclusion not present/appropriate	O Loosely organized O Relationship between ideas sometimes unclear O Several problems with cohesion, sequencing, and flow of ideas	G Basic sentence variety G Two or three sentence combination problems G Several errors in phrase structure/parallelism G Frequent errors in grammatical forms G Several missing constituents or function words V Adequate basic vocabulary but lacks sophistication and range. V Repetition of basic vocabulary V Many errors (7+) in word form, word choice, or idiomatic phrasing V Meaning is sometimes unclear	M Frequent errors in spelling, punctuation, and/or capitalization (7+); errors make meaning unclear F Problems with paragraph form F Messy presentation
4	Below standard C No clear central theme C Poor development of topic C Support is mostly vague, insufficient, unconvincing, and/or off-topic Not enough to evaluate	O Ideas not organized O Relationship between ideas often unclear O Difficult to follow	G Absence of complex sentences G Several sentence combination problems G Problems with simple sentences G Frequent errors in phrase structure/parallelism G Frequent errors in noun or verb forms G Many missing constituents or function words V Problems with basic vocabulary V Frequent errors in word choice and word form V Very narrow range of vocabulary V Meaning is often unclear	M Poor mastery of two or more of the following: spelling, punctuation, and/or capitalization; errors often make meaning unclear F Writing is illegible

Legend: **C** = Content; **O** = Organization; **G** = Grammar; **V** = Vocabulary; **F** = Form; and **M** = Mechanics.

C: ___/10 O: ___/10 GV: ___/10 FM: ___/5

Appendix B. “Dreams never Die”: Prewriting task

Writing Topic:

You will write a paragraph about yourself or someone you know personally (i.e., **NOT** a famous person) and describe how you/this person overcame an obstacle. In your paragraph, you should . . .

- name the obstacle you/this person faced and describe/explain it a little bit.
- explain how you/this person overcame the obstacle (this is the **main focus** of the paragraph). If applicable, you may want to mention what benefits you/this person obtained as a result.

Part 1: Generating & Evaluating Ideas

- (a) Review the list of obstacles collected on the blackboard, and think of your own life and of the people you know personally (family, friends, classmates, colleagues, . . .). Who has faced an obstacle in their life? How did these people overcome that obstacle? Write down the following information in the table below: the person’s name, the obstacle they faced, and how they overcame that obstacle.

Name	Obstacle	How overcame the obstacle	Any benefits	Appropriate to write about?

- (b) Tell your group about the people you thought of, and what obstacle they faced.
- (c) As you listen to your group talking about the people they know, think about whether it is appropriate to write about them for this assignment. Tell your group why you think each person would or would not be appropriate for this assignment.

Part 2: Selecting & Organizing Ideas

- (a) Considering the feedback of your group, choose the best person to write about. What information will you mention in which order? Make an outline, and then share it with your group.
- -
 -
 -
- (b) As you listen to your group’s writing plans, give them feedback about whether their outline is well-organized.

References

- Abadikhah, S., & Mosleh, Z. (2011). EFL learners’ proficiency level and attention to linguistic features during collaborative output activities. *Iranian EFL Journal*, 7, 179–198.
- Alegría de la Colina, A., & García Mayo, M. d. P. (2007). Attention to form across collaborative tasks by low-proficiency learners in an EFL setting. In M. d. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 91–116). Clevedon, England: Multilingual Matters.
- DiCamilla, F. J., & Anton, M. (1997). Repetition in the collaborative discourse of L2 learners: A Vygotskian perspective. *Canadian Modern Language Review*, 53, 609–633.
- Donato, R. (1988). *Beyond group: A psycholinguistic rationale for collective activity in second language learning*. (Unpublished doctoral dissertation). Newark: University of Delaware. Retrieved from: <http://search.proquest.com/docview/303711939>
- Donato, R. (1994a). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.

- Donato, R. (1994b). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 27–50). Oxford: Oxford University Press.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30, 474–509. <http://dx.doi.org/10.1093/applin/amp042>
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59–84. <http://dx.doi.org/10.1017/S0272263104026130>
- Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning and Technology*, 14(3), 51–71.
- 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. <http://dx.doi.org/10.1016/j.jslw.2011.12.002>
- Fernández Dobao, A. (2014). Attention to form in collaborative writing tasks: Comparing pair and small group interaction. *Canadian Modern Language Review. La revue canadienne des langues vivantes*, 70(2), 158–187.
- Fernández Dobao, A., & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, 41(2). <http://dx.doi.org/10.1016/j.system.2013.02.002>
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323. <http://dx.doi.org/10.1017/S0272263100015047>
- García Mayo, M. d. P. (2002). The effectiveness of two form-focused tasks in advanced EFL pedagogy. *International Journal of Applied Linguistics*, 12, 156–175.
- Gutiérrez, X. (2008). What does metalinguistic activity in learners' interaction during a collaborative L2 writing task look like? *Modern Language Journal*, 92, 519–537. <http://dx.doi.org/10.1111/j.1540-4781.2008.00785.x>
- Higgins, L., Flower, L., & Petraglia, J. (1992). Planning text together. *Written Communication*, 9(1), 48–84. <http://dx.doi.org/10.1177/0741088392009001002>
- Ismail, L., & Samad, A. A. (2010). The effect of tasks on language-related episodes (LREs) during focus-on-form instruction. *Language Education in Asia*, 1, 87–98.
- Kagan, S. (1995). Group grades miss the mark. *Educational Leadership*, 52(8), 68–71.
- Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal*, 92(1), 114–130. <http://dx.doi.org/10.1111/j.1540-4781.2008.00690.x>
- Kim, Y. (2009). The effects of task complexity on learner–learner interaction. *System*, 37, 254–268.
- Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12, 211–234. <http://dx.doi.org/10.1177/1362168807086288>
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183–199. <http://dx.doi.org/10.1177/1362168810388711>
- Kroll, B. (1990). What does time buy? ESL student performance on home versus class composition. In B. Kroll (Ed.), *Second language writing research: Insights for the classroom* (pp. 140–154). Cambridge: Cambridge University Press.
- Kuiken, F., & Vedder, I. (2002). Collaborative writing in L2: The effect of group interaction on text quality. In S. Randsdell M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 169–188). Dordrecht/Boston/London: Kluwer.
- Kuiken, F., & Vedder, I. (2005). Noticing and the role of interaction in promoting language learning. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 353–381). Berlin: Mouton de Gruyter.
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8, 55–81.
- Manchón, R. M., & Roca de Larios, J. (2007). On the temporal nature of planning in L1 and L2 composing. *Language Learning*, 57(4), 549–593. <http://dx.doi.org/10.1111/j.1467-9922.2007.00428.x>
- McDonough, K., Neumann, H., & Trofimovich, P. (2014). Eliciting production of L2 target structures through collaborative priming activities. *Canadian Modern Language Review*.
- Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System*, 34, 566–585. <http://dx.doi.org/10.1016/j.system.2006.08.003>
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–148.
- Roberts, T. S. (2006). Self, peer and group assessment in E-Learning: An introduction. In T. S. Roberts (Ed.), *Self, peer and group assessment in E-Learning*. Hershey, PA: Information Science Publishing.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20, 286–305. <http://dx.doi.org/10.1016/j.jslw.2011.05.010>
- Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *Journal of Second Language Writing*, 7, 319–345. [http://dx.doi.org/10.1016/s1060-3743\(98\)90020-0](http://dx.doi.org/10.1016/s1060-3743(98)90020-0)
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). Cambridge: Cambridge University Press.
- Storch, N. (2002a). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158. <http://dx.doi.org/10.1111/1467-9922.00179>
- Storch, N. (2002b). Relationships formed in dyadic interaction and opportunity for learning. *International Journal of Educational Research*, 37, 305–322. [http://dx.doi.org/10.1016/s0883-0355\(03\)00007-7](http://dx.doi.org/10.1016/s0883-0355(03)00007-7)
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14, 153–173. <http://dx.doi.org/10.1016/j.jslw.2005.05.002>
- Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research*, 11, 143–159. <http://dx.doi.org/10.1177/1362168807074600>

- Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness*, 17(2), 95–114. <http://dx.doi.org/10.1080/09658410802146644>
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters.
- Storch, N., & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research*, 17, 31–48. <http://dx.doi.org/10.1177/1362168812457530>
- Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. d. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Clevedon, England: Multilingual Matters.
- Strauss, P., & U, A. (2007). Group assessments: Dilemmas facing lecturers in multicultural tertiary classrooms. *Higher Education Research & Development*, 26(2), 147–161. <http://dx.doi.org/10.1080/07294360701310789>
- Suzuki, W., & Itagaki, N. (2007). Learner metalinguistic reflections following output-oriented and reflective activities. *Language Awareness*, 16(2), 131–146. <http://dx.doi.org/10.2167/la392.0>
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320–337.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 99–118). Harlow, England: Longman.
- Sweigart, W. (1991). Classroom talk, knowledge development, and writing. *Research in the Teaching of English*, 25, 469–496.
- Trofimovich, P., McDonough, K., & Neumann, H. (2013). Using collaborative tasks to elicit auditory and structural priming. *TESOL Quarterly*, 47(1), 177–186. <http://dx.doi.org/10.1002/tesq.78>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121–142. <http://dx.doi.org/10.1177/136216880607074599>
- Watanabe, Y., & Swain, M. (2008). Perception of learner proficiency: Its impact on the interaction between an ESL learner and her higher and lower proficiency partners. *Language Awareness*, 17(2), 115–130. <http://dx.doi.org/10.1080/09658410802146651>
- Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14(1), 85–106. <http://dx.doi.org/10.1177/026553229701400105>
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26, 445–466. <http://dx.doi.org/10.1177/0265532209104670>
- Wigglesworth, G., & Storch, N. (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing*, 21, 364–374. <http://dx.doi.org/10.1016/j.jslw.2012.09.005>

Heike Neumann is a lecturer of English as a second language and a language test developer in the Department of Education at Concordia University. Her research interests include second-language writing pedagogy, writing assessment, and language assessment.

Kim McDonough is an associate professor and Canada Research Chair in applied linguistics in the Department of Education at Concordia University. Her research interests include interaction, task-based language teaching, and psycholinguistic approaches to second language acquisition.