English for Specific Purposes 62 (2021) 128–141

Contents lists available at ScienceDirect

English for Specific Purposes

journal homepage: http://ees.elsevier.com/esp/default.asp

An exploratory analysis of source integration in post-secondary L1 and L2 source-based writing

Stephen M. Doolan

Texas A&M University – Corpus Christi, 6300 Ocean Dr., Corpus Christi, TX 78412 USA

ARTICLE INFO

Article history:
Available online 16 February 2021

Keywords:
Reading-to-write
Source-based writing
Mixed-methods
L2 writing
First-year composition

1. Introduction

In the US, First-Year Composition (FYC) courses are typically taken during the first semester of post-secondary studies and they often function to prepare students for writing across the curriculum. As such, FYC classes are most accurately described as English for General Academic Purposes, rather than the English for Specific Academic Purpose classes we would find nested within disciplinary contexts (Basturkmen, 2017). Yet, within this wide-angled, generic context for writing instruction, building FYC curriculum around source-based writing is appropriate for a number of reasons. First, integration of source material within student writing is a cognitively demanding task that requires use of various types of source integration (e.g., citation, paraphrasing, quoting) and selection of the source material that most effectively supports the writer's purposes. Furthermore, it is an obligatory form of writing throughout much of post-secondary education (Carson, 2001; Rosenfeld, Leung, & Oltman, 2001). Thus, students benefit from explicit instruction on source-based writing (Hirvela, 2016). Source-based writing is also an appropriate FYC focus because evidence exists that well-designed writing courses for second language (L2) students can build writing skills that transfer to stronger disciplinary writing in students' future fields of study (James, 2018). While integrating source material in student writing is challenging in a first language (L1), it is obviously made all the more difficult in a second language, where L2 learners must navigate a host of additional language-related issues

E-mail address: Stephen.Doolan@tamucc.edu.

https://doi.org/10.1016/j.esp.2021.01.003
0889-4906/

(Grabe & Zhang, 2013). To improve the quality of L1 and L2 source-based writing instruction, more empirical research into source-based writing is needed (Cumming, Lai, & Cho, 2016).

This article presents an exploratory analysis of source text use within L1 and L2 student writing. Use of source material in student writing is studied through an investigation of various *types of source integration* and *ideational units*. Analysis of the types of source integration includes identification of both citation practices (e.g., integral, non-integral, and missing citation) as well as how the source content is presented: either as a paraphrase or a quotation. Ideational units are specific ideas or themes appropriated from the source text and incorporated into students' written sentences. Using survey information to distinguish student groups and timed, in-class writings on two problem-solution prompts, this study draws on a relatively large sample of student writing ($N = 149$) to compare source text use in a mixed-methods, exploratory analysis. This study provides a meaningful contribution to reading-to-write research by conducting a thorough, descriptive, discourse analysis of L1 and L2 student writing from multiple source texts at what is commonly the beginning of these students' post-secondary studies.

2. Literature review

Reading-to-write research has been a productive area of inquiry, with carefully controlled studies spanning over three decades (see Hirvela, 2016 for an overview of existing research). Cumming et al. (2016) conducted a synthesis of 69 source-based empirical writing studies between the years of 1993 and 2013 and were able to identify a number of overarching themes including the fact that meaningful differences exist between L1 and L2 writing as they relate to integrating sources. Another large scale synthesis by Liu, Lin, Kou, and Wang (2016) identified, among other findings, the ongoing need for L2 writer development in the areas of integrating source use and attributing sources appropriately. Grabe and Zhang (2013) highlighted specific reading-to-write challenges for L2 writers such as (a) L2 reading and L2 writing proficiency, (b) fluency issues in both L2 reading and L2 writing, (c) lack of experience with L2 reading-to-write tasks, (d) more limited L2 vocabulary compared to L1 students, and (e) more limited background knowledge and L2 cultural knowledge to aid in the L2 reading/writing process. These three studies indicate that source-based writing challenges faced by L1 and L2 students are different and that a greater understanding of source integration processes in L1 and L2 writing respectively would have pedagogical and research value. Moving from the general challenges discussed above, the remainder of this review will focus on specific source-based writing issues, namely source-based writing scholarship as it relates to writing quality, types of source integration, and the integration of ideas drawn from source texts.

2.1. Source-based writing and holistic writing quality

Research on L1 source-based writing and holistic writing quality highlights the importance of L1 reading ability and identifies specific challenges with less successful L1 source-based writing. As discussed by Hirvela (2016), Stotsky (1983) was the first L1 scholar to place considerable emphasis on source-based writing variables and writing quality. Stotsky reported that better readers were "almost consistently" better writers, and better writers "tend to read more than poorer writers" (p. 636). In a study of 71 college undergraduates, Risemberg (1996) also found that reading ability was a unique contributor to writing quality. Finally, research by Spivey and King (1989) investigated various levels of K-12 synthesis writing abilities and concluded that stronger reading ability led to more source use, better targeting of important information drawn from the source text, better local and global coherence in the student writing, and a higher holistic writing quality. Other challenges tied to lower holistic quality, L1, source-based writing include less effective summarization and transformation of source concepts (Howard, Serviss, & Rodrigue, 2010; McGinley, 1992; Sole, Miras, Castells, Espino, & Minguela, 2013). The above research suggests that there are undoubtedly relationships between reading ability, the holistic quality of L1 source-based writing, and the effectiveness of source use.

Studies in L2 writing research have found that holistic writing quality has a strong relationship to more and less effective source use. Scholars have suggested that manipulating, summarizing, and transforming source material is a challenging task for L2 student writers (Leki & Carson, 1997; Plakans, 2009a; Zhao & Hirvela, 2015). Cumming et al. (2006) analyzed L2 writing on independent and integrated writing tasks and identified different source use patterns at high, mid, and low holistic writing quality scoring bands. The L2 writing that earned high holistic writing quality scores contained more coherent summaries and syntheses and these writers wrote with more fluency, findings confirmed in other empirical studies (Gebril & Plakans, 2009, 2013). The mid proficiency level was marked by some paraphrasing as well as "verbatim piecemeal phrases from sources" (p. 45). Finally, the writers who received lower holistic writing quality scores may not have been skilled enough to write from sources, instead relying on verbatim phrases from source texts and generally less use of paraphrasing and summarizing compared to the writers who achieved mid and high holistic writing quality scores. These verbatim phrase findings at the mid and low holistic writing quality levels connect with a large body of empirical research linking L2 writing and plagiarism (Campbell, 1990; Currie, 1998; Johns & Mayes, 1990; Keck, 2014). Finally, studies by Asención Delaney (2008), Kim (2001), Plakans and Gebril (2012), Sun (2012), and Weigle (2004) suggest that summarization from more or less challenging texts, and challenges with reading comprehension affect writing quality for L2 writers.

The above research suggests that L1 and L2 students with different holistic writing quality levels are likely to integrate source material into their writing more and less successfully. Reporting holistic writing quality in the current study is an attempt to acknowledge an important overarching construct that clearly influences source use.

2.2. Types of source integration

Previous empirical research on student writing has provided contradictory evidence in terms of how frequently various types of source integration (e.g., paraphrases, quotations) are employed by L1 and L2 student writers. Two studies suggest relatively high frequencies of quotations compared to paraphrases in student writing. Recent work by Jamieson (2013) investigated 174 essays from FYC students at 16 US colleges and universities to better understand source use among L1 student writers. Analyzing 1,911 student citations, Jamieson found that quotations were the most frequent type of source use, accounting for 41.5% of all citations, followed by paraphrases at 31.9%. Comparative research suggests that L2 writers quote more than L1 writers (Campbell, 1990). Yet, other empirical studies have shown a strong student preference for paraphrasing or summarizing over quotation. Lee, Hitchcock, and Casal (2018) investigated 100 research papers from L2 FYC students at Ohio University and found that L2 students quoted approximately one time per 1,000 words and used some form of summary, paraphrase, or synthesis approximately 8.48 times per 1,000 words on first drafts. Finally, Storch (2012) investigated primarily Chinese L1 undergraduate students in an economics/commerce writing class and confirmed the relatively smaller number of quotations in L2 student writing relative to paraphrases and summaries.

The findings presented above are inconsistent in terms of students' frequencies of quotations relative to paraphrases. Part of this discrepancy may be due to methodological differences. Studies such as Jamieson (2013) analyzed only source use that included a citation (i.e., instances where citations were provided), while other studies analyzed source texts and student texts to make independent determinations of when sources had been used. Because quotations are more likely to be cited, analyzing only source use that includes citations is likely to inflate the ratio of quotation to paraphrase. Of course, relying less on citations and more on independent determinations of whether a source was used involves more complicated analysis and controversial decisions around plagiarism and students' intent.

Studies of the types of source integration in L1 and L2 student writing have found that citations are often entirely missing in quite a few instances where they would be expected. In Storch (2012), 26 L2 writers at a high intermediate level (IELTS scores at approximately 6.5) completed timed, in-class writings using a pre/post design. In the first writing task, Storch found 205 paraphrases, none of which were cited. In comparing L1 and L2 writing, work by Campbell (1990) found that L2 students provided citations for quotes and paraphrases 42% of the time, compared to only 16% of the time for L1 students. To understand why students may or may not be including citations, Shi (2008) conducted text analysis and interviews with 16 undergraduates from seven different L1 backgrounds (three students with L1 English) examining writing projects in mostly lower-level university courses. Findings indicated that, when interpreting source text, students did not cite sources 35% of the time. The students provided a host of reasons for why they did not cite sources (e.g., the information was common knowledge, the reference was cited earlier in the text). This research suggests that whether or not to cite source material represents a judgment that "… is negotiated, localized, and contingent" (p. 22). To provide one example from Shi (2008), students new to post-secondary written academic source use must determine the potentially blurry and subjective lines between common knowledge (no citation needed) and appropriated knowledge (citation needed). Because students are still learning these practices, there is a certain amount of trial and error, as students see examples of source use within academia (with or without citations), use sources themselves, and receive feedback from instructors within and across disciplines on appropriate and inappropriate source use. In this way learning to cite sources is negotiated with the instructor, but also within disciplines and more local institutional settings. Based on interviews from Shi's study, students may be making conscious decisions not to cite sources as they develop academic literacy skills.

In summary, existing research has found conflicting patterns of frequency in citing, quoting, and paraphrasing in L1 and L2 student writing. There is also some confusion among both students and teachers as to what content needs to be cited.

2.3. Source-based writing challenges and idea units

As students compose using a source-based writing prompt, the source ideas are central to the paraphrasing, synthesizing, and transforming of source material taking place (Hirvela, 2016). Surprisingly, very few empirical studies have explored and compared which source ideas are used by L1 and L2 student writers. Teachers often encourage students to focus on the main ideas (Dollahite & Haun, 2012), which could lead to the hypothesis that a small number of the most important ideas will account for the majority of the source ideas appropriated into successful source-based writing, though this researcher is unaware of any empirical research to support this hypothesis. Existing research suggests that students use the location of an idea in the source texts rather than its importance as criteria for inclusion in their own writing. For example, Gradwohl Nash, Schumacher, and Carlson (1993) suggest that students are more like to use ideas from the first source when multiple texts are provided, and Jamieson (2013) found that 82.62% of all cited source use was drawn from the first four pages of source texts, regardless of the source text's length.

The above literature review has focused on the quality of L1 and L2 source-based writing, the various types of source integration, and the ideational units used in student writing. Yet, while the above studies certainly confirm the challenges of source-based writing tasks for post-secondary students, there are still too few studies (1) comparing how L1 and L2 holistic writing quality at the FYC level may be interacting with source-based writing, (2) exploring the frequencies of different types of source integration among L1 and L2 writers, and (3) analyzing the range of ideas that student writers are using from the source text. The current study explores these areas under controlled conditions where potentially mediating variables such as time-on-task, distractions in the environment, and assistance from friends and family are less influential. More research in

these areas will help pinpoint the strengths and limitations that students bring to the FYC classroom as they look to further develop their source-based writing skills. As such, the current study fills a gap by attending to this complex host of textual and contextual issues and is guided by the following research questions.

2.4. Research questions

The current study begins by assigning holistic writing quality scores in order to establish a basic measure of writing proficiency. With a general understanding of writing proficiency in place, the current study asks the following research questions:

1. What patterns emerge in terms of the types of source integration used by FYC L1 and L2 students on a 40-min, problem-solution, source-based writing task?
2. What patterns emerge in terms of which ideational units from the source texts are used by FYC L1 and L2 students on a 40-min, problem-solution, source-based writing task?

3. Method

3.1. Participants and setting

Both the L1 ($n = 112$) and L2 ($n = 37$) participants were drawn from one medium-sized, public, regional institution in South Texas. Participants were enrolled in 13 sections of First-Year Composition (FYC) courses. In U.S. educational contexts, FYC is typically taken very early in students' post-secondary studies, and the FYC course is seen as a general introduction to writing in post-secondary contexts. The L1 students indicated that, with their families (or in the house where they spent most of their childhoods), they did not regularly speak a language other than English. Of the L1 students, 96% were 18–19 years old. Students who spent extensive time in the U.S. educational system but had grown up regularly speaking a language other than English at home (i.e., resident L2 students) were excluded from the current study because legitimate uncertainty exists regarding whether these students more closely resemble L1 or L2 writers (Doolan, 2017) and exploration of this issue is beyond the scope of the current study.

The L2 students came from five different L1 backgrounds, with the largest groups being L1 Chinese ($n = 14$, or 38% of total L2 students), L1 Vietnamese ($n = 12$, or 32% of total L2 students), and L1 Arabic ($n = 6$, 16% of total L2 students). There were also four L1 Spanish speakers and one L1 Korean student. Of the L2 students, 41% were 18 or 19 years old. The remainder of the L2 students were between the ages of 20 and 32. None of the L2 students had studied in the U.S. for more than 2 years. They were all studying in the U.S. on international visas, a further indication that these were international L2 and not resident L2 students.

3.2. Instruments and materials

Instructions (including a script) were distributed to the instructors administering this project. This document provided guidance for how to prepare and administer the project. The second key instrument used in the current study was a self-contained online module. The module was comprised of an informed consent request, a demographic survey to help determine the student group, and finally the source-based writing task. The demographic survey guided L1 and L2 students to different sets of questions and took approximately 5-min to complete. The survey questions confirmed eligibility for inclusion (e.g., age), and group inclusion (i.e., criteria for L1 and L2 categories).

The two problem/solution prompts with the four accompanying source texts (two per prompt) were used (with permission) from Zhang (2013).[1] The source texts were tested and adapted to ensure that the vocabulary level, Flesh-Kincaid grade level, and reading length were comparable across readings and appropriate for advanced language learners (see Zhang, 2013, p. 56 for details). Additionally, the topics (i.e., Prompt 1 – Study Abroad, Prompt 2 – Global Warming) were selected to increase the likelihood that both L1 and L2 students would be familiar with the text topics.

3.3. Procedures

Data were collected in the first two weeks of instruction during two separate semesters: Fall 2013 and Fall 2015.[2] Seven different instructors were provided scripts and instructions for administration of this project in their FYC courses. Using the scripts in computer labs on campus, instructors introduced the research project and guided students to the online module.

[1] See Zhang (2013) for full source texts or the supplemental material found at the *English for Specific Purposes Journal* website.

[2] Recruitment procedures and materials (including instructions and a script for administration) were identical in 2013 and 2015. Furthermore, comparable populations were used (i.e., same institution, same course) in both rounds of data collection. All of the L1 and 12 of the L2 essays were collected in 2013 with the remaining 25 L2 essays collected in 2015.

Students were given approximately 40-min to complete the source-based writing task (the maximum time possible to fit within a 75-min class), and were randomly assigned to one of two possible prompts (with associated readings). Students were instructed to compose their essays in a Word document and cut and paste the document into the text box (within the module) at the end of the allotted time. This task was also used as a beginning-of-semester classroom diagnostic. As such, this assignment served as both a writing sample for this research, and also as an authentic, in-class writing assignment. Fidelity of administration emails were sent to all participating instructors/administrators, confirming that in-class writing conditions were comparable across the different data collections.

3.4. Analysis

From the 13 classes participating in the study, approximately 260 essays were collected. All essays from this larger sample written by students who fit the L2 student criteria were included. This was because there were many fewer L2 students enrolled in FYC courses at the target institution compared to L1 students. Of the L1 essays, the researcher employed stratified random sampling to ensure approximately equal numbers of attempts on the two writing prompts. The researcher used as large an L1 sample as was feasible, and while this resulted in more L1 than L2 participants, uneven sample sizes are not uncommon in mixed-method writing research (Polio & Friedman, 2016, p. 89), allowing for the most robust quantitative findings as possible under the circumstances. The word count for the L2 writing ($n = 37$) was 12,207, and the L1 writing ($n = 112$) totaled 51,884 words.

To determine holistic quality of student writing, four FYC instructors conducted three rounds of calibration for assigning scores on the five-point rubric used by Zhang (2013) (Appendix A). Following calibration, each essay received two independent scores with the scores summed, resulting in holistic quality scores ranging from zero to 10. If scores between raters differed by more than one scoring band (e.g., rater one assigns a three and rater two assigns a five), a third rater was used to systematically adjudicate the final score assigned. Intraclass correlation coefficients based on a two-way random effects model was used to measure interrater reliability, resulting in an ICC of .87 (good reliability) on all essays scored (with 94% of the essays within one scoring band).

To determine whether holistic quality scores of L1 students differed significantly from the scores received by L2 students, an independent samples t-test was conducted. All assumptions were met, including the assumption that these data were normally distributed, and the statistical analysis indicated a significant main effect, $t(150) = 4.821$, $p < .05$, with a Cohen's d effect size of .9224 (large) and $r = .419$. This t-test provides statistical support for the claim that the holistic writing quality of the L1 writing ($M = 6.14$, $SD = 2.134$) received significantly higher scores than the writing of the L2 students ($M = 4.33$, $SD = 1.774$). While discussions of holistic quality do not factor heavily into the results and discussion sections, evidence of differing levels of L1 and L2 holistic writing quality is relevant when comparing the current study to previous (Asención Delaney, 2008; Campbell, 1990; Cumming et al., 2006, 2016; Plakans & Gebril, 2012) and future research.

Both research questions in the current study investigate the larger construct of integrating source material into student writing. This larger construct can be split into two components that align with the two research questions: (1) *how* students are integrating source text material, or their types of source integration, and (2) *what* information, or ideational units, students are appropriating from the source texts.

The analysis of types of source integration and ideational units involved the creation of two coding schemes (Tables 2 and 3[3]). The chronological series of steps for creating and completing these analyses are presented below.

- The two coding schemes (RQ1, RQ2) were created by the researcher.
- The researcher and two FYC instructors applied the coding schemes to eight student essays, re-visioning and refining the coding schemes.
- The previous step was repeated with another eight student essays.
- Interrater reliability was calculated on 17 new essays. On these essays, only coding decisions with agreement between the raters were used for further analysis. Following this final round of coding, systematic coding descripancies were again discussed to further improve reliability.
- The initial 16 calibration essays were revisited, with the lead researcher and the two coders applying the finalized coding scheme and agreeing on final coding decisions.
- The remaining 116 essays were divided between the two raters, and each rater coded approximately half of the remaining essays.
- After all essays were coded, the computer program Monoconc was used to gather frequency counts (normalized per 1,000 words) for all codes.[4]

[3] While these two Tables provide the overarching framework for the discourse analysis, a series of additional decisions were made in order to ensure systematic coding throughout the process. Readers may contact the author for more details about these additional decisions.

[4] Counts per text would have resulted in too many cells with zero occurrences, making for less meaningful (and less robust) inferential statistical results. As such, the researcher followed the type of descriptive corpus-based research methodology that we see frequently in exploratory corpus linguistic work and presents normed frequency counts.

- Finally, Monoconc was used to identify codes with interesting quantitative patterns of difference between L1 and L2 writing for additional qualitative analysis.
- For codes analyzed qualitatively, all sentences given the target code were analyzed, as well as surrounding sentences in an effort to determine functional patterns.

3.4.1. Coding scheme details

Types of source integration codes (RQ1) and ideational codes (RQ2) were assigned at the sentence level rather than to clauses or T-units. The decision to establish the unit of analysis as the sentence was made to ensure greater reliability of coding. Though both coders had advanced degrees in English, they had limited linguistics training and using the clause, T-unit, or idea-unit from the source would have introduced potentially inconsistent unit boundaries across coding decisions, possibly compromising reliability (see Polio & Friedman, 2016, p. 111 for examples of these unit boundary complications). Coders were instructed to assign all codes that applied to a given sentence, meaning that multiple types of source integration codes and/or multiple ideational codes could be assigned in any given sentence. Allowing multiple codes per sentence further mitigates the potential to under-report codes in each sentence. The concept of ideational codes may be loosely tied to *idea* units discussed in Plakans (2009a) and Johns and Mayes (1990), but was not modeled after any previously employed reading-to-write coding system.

As shown in Table 1, the categories for types of source integration are likely familiar to scholars who have worked with citation practices. That said, some categories involve more clear-cut decisions than others. For example, quotations can largely be distinguished with the presences of quotation marks. On the other hand, categories involving paraphrasing require more subjective determinations of whether ideas were drawn from the source text or from the writers' preexisting familiarity with the topic. In terms of paraphrases, a code was assigned to a target sentence, and thus considered a paraphrase if direct links could be made between the target sentence in student writing and ideational units identified in the source texts. The direct links were determined by comparing propositional material within the target sentences and the propositional material within the source-based coding categories presented in Table 2.

As shown in Table 2, a total of 15 ideational units were identified in the two sets of source texts (i.e., four readings). With all four readings totaling approximately 1,800 words, it is worth noting that 15 ideational units is a relatively small number. Often, those units aligned with major topics discussed in given paragraphs of the source texts.

3.4.2. Interrater reliability

Interrater reliability for each coding scheme was measured using precision by first identifying sentences where both raters provided one or more coding decision and then calculating how precisely the coders agreed on the code(s) assigned. Measuring interrater reliability involved the use of three contingency tables, one for the types of source integration (Kappa = .935), one for ideational units from Prompt 1 (Kappa = .927), and a final table for ideational units for Prompt 2 (Kappa = .907).

Table 1
Categories for coding the types of source integration.

Code title & abbreviations	Code descriptions[a]
Direct quotation attributed (to a source) [dqa]	Either in a parenthetical at the end of the sentence or within the writer's sentence (e.g., *According to Smith …*), the writer "points" to the text in the target sentence.
Direct quotation indirectly attributed [dqia]	Though the sentence under consideration does not include *any* direct mention of the source text, there is mention of the source text elsewhere in the essay.
Direct quotation unattributed [dqua]	While a quote is present, there is no mention of the source text in the target sentence nor is there mention of the source text elsewhere in the student essay.
Paraphrase, (source) referenced in text [prt]	This type of paraphrase includes an in-text citation such as *According to* or *As stated by* within the target sentence where the paraphrase occurs. This reference can be made before the paraphrase, during, or after, but should occur within the target sentence.
Paraphrase, (source) referenced parenthetically [prp]	A parenthetical (e.g., *(Smith, 2015)*) is connected to the sentence in which the paraphrase occurs.
Paraphrased, (source) indirectly referenced [pir]	Though the sentence under consideration does not include *any* direct mention of the source text (neither in text nor parenthetically), it is referenced (either in text OR parenthetically) elsewhere in the essay. While such practices can certainly be seen as inappropriate, "illegitimate intertextuality," and plagiarism, Pecorari and Shaw (2012) suggest that such missing citations *may* appear to other instructors as not plagiarism, and further, as "not only appropriate, but productive and beneficial, both to the learning process and to the text" (p. 159).
Paraphrased, no (source) reference [pnr]	There is an ideational link to the source text (whether the writer meant to make that link or not), but no mention of the source is anywhere in the essay.
Plagiarism possible [above codes + pp]	Plagiarism, in this rather impressionistic category, can be considered the use of three or more identical words taken directly from the source text at one stretch without punctuation to indicate the language is coming from the source. Also, plagiarism could be "patchwriting" in which the same sentence structure is present with a word occasionally changed or added. It is also worth noting that this category is secondary to various types of quotes or types of paraphrases. In short, does the sentence *feel* plagiarized?

[a] See supplemental material on the *Journal of English for Specific Purposes* website for the full code descriptions.

Table 2
Categories for coding of ideational units.

Source text	Code & abbreviations	Select (source-based) themes[a]
Prompt 1, Text 1 (Asino)	Culture shock [cs]	• Loss of emotional balance • Emotional illness may ensue
	Difficulty making friends [dmf]	• Friendly does not equal a commitment to friendship • American students value personal freedom … less open to others outside of comfort zone
	Classroom Issues [ci]	• Instructors don't notice students • Topics not international friendly … but everyone needs to participate
Prompt 1, Text 2 (Coomer)	Pre-departure strategies (for culture shock) [pds]	• Expectations affect experience • Read pre-departure materials that schools send • Ask friends who have studied in U.S.
	Post-arrival strategies (for culture shock) [pas]	• (after advice on personal safety) explore environment • Be courageous • Keep in touch with own culture … will reduce homesickness
	Holistically shape [hs]	• Well-nourished physically, mentally, emotionally, and socially • … even if over-busy from school
Prompt 2, Text 1 (NRDC)	Heating related [hr]	• 1998–2007, among 25 hottest in US • Could warm 7.2° in 21st century
	Weather effects [we]	• More powerful hurricanes, more droughts & wildfire • Hurricanes: warmer water, more energy, stronger storms
	Geological effects [ge]	• Glaciers & Sea levels • Artic summers ice-free by 2040
	Human health effects [hhe]	• Heat waves & air pollution • Hotter weather equals more air quality problems … already affecting 100 million Americans
Prompt 2, Text 2 (Radford)	Renewable energy sources [res]	• Solar energy – it doesn't use dirty energy
	Recycling Issues [ri]	• Recycling aluminum, glass, & plastic, can save half of energy needed to make from scratch • Recycled paper reduces # of trees cut … allows trees to do job (CO_2 reduction, O_2 creation)
	Home Conservation [hc]	• Wait until evening to run appliances • Reduce water use: low flow faucets & toilets … fixing dripping
	Vehicle related [vr]	• Use a bike, public trans, walk • Hybrid vehicles (the price has come down)
	Reducing wasteful usage [rwu]	• Dirty energy takes away precious natural resources

[a] See supplemental material on the *Journal of English for Specific Purposes* website for the full list of themes.

4. Results

4.1. Research question 1: What patterns emerge in terms of the types of source integration used by FYC L1 and L2 students on a 40-min, problem-solution, source-based writing task?

Table 3 shows the normalized frequency counts for L1 and L2 texts in terms of the types of source integration (e.g., types of citations, paraphrases, and quotations). As shown in Table 3, the use of quotations for both groups was minimal, with all forms of quotation occurring slightly more than twice per 1,000 words. When quoting did occur, 79% of quotes (both L1 and L2) were directly attributed to a source (e.g., *She says that culture shock is a "documented psychological occurrence"*). In contrast, all forms of paraphrasing were much more likely to occur in both the L1 and L2 student texts (L1 = 20.53, L2 = 21.38) when compared to all forms of quotation (L1 = 3.03, L2 = 2.21). Unlike the quotations, when students included some form of paraphrase, they were much *less* likely to attribute the paraphrase in the target sentence to the source that was being used. The directly attributed paraphrasing accounted for approximately 8% of the L1 paraphrases, and 15% of the L2 paraphrases (e.g., *By Tuna Asino and Patrick Coomer's reaserching, the most 2 important problems are culture shock and making friends.*).

Table 3
Normalized frequency counts (per 1,000 words) for types of source integration.

Reference type	L1 texts	L2 texts
Direct quote attributed to source (dqa)	2.35	1.80
Direct quote indirectly attributed to source (dqia)	0.23	0.25
Direct quote unattributed (dqua)	0.44	0.16
Paraphrase, reference in text (prt)	1.41	2.21
Paraphrase, referenced parenthetically (prp)	0.25	0.98
Paraphrase, indirectly referenced (pir)	9.52	12.12
Paraphrased, no reference (pnr)	9.35	6.06
Plagiarism possible (pp)	0.29	0.16

Comparison of L1 and L2 type of source integration patterns revealed that quotations were used slightly less frequently by L2 students compared to L1 writers (L2 = 2.21, L1 = 3.03 for all quotes). Analysis also found that paraphrasing with reference to the text (L2 = 3.19, L1 = 1.66) as well as paraphrasing with indirect referencing (L2 = 12.12, L1 = 9.52) (i.e., the reference occurring elsewhere in the student text) was more frequent in L2 writing than in L1 writing. Finally, compared to L2 writers, L1 students produced 54% more paraphrases with no reference to the source text anywhere in the students' text.

The impressionistic coding of possible plagiarism resulted in very low numbers for both groups (less than .30 per 1,000 words). Whether this number adequately accounts for what would consistently be considered plagiarism by college instructors is a complicated question and is beyond the scope of this manuscript (Pecorari & Shaw, 2012). Instead, this number simply provides us with the amount of text that was *perceived* by these coders to be identical or nearly identical source text language use without attribution or quotation marks.

In sum, both student groups used considerably more paraphrasing than quotation, and were much more likely to directly cite quotations within a target sentence than to indirectly or not attribute a quote to a source citation. Both groups were much less likely to cite paraphrases within the target sentences than to indirectly or not attribute paraphrases to a source citation. Also, L2 students were slightly less likely to directly attribute quotations to source text in target sentences than L1 students, but more likely than L1 students to reference "responsibly" when it came to paraphrasing.

Because the category of *paraphrase, indirectly referenced (pir)* was the most frequently occurring category in Table 3 (accounting for 46% of all types of source integration across the whole sample), it was selected as a meaningful category for follow-up qualitative analysis. Indirect paraphrases were operationalized as sentences that do not include *any* direct mention of the source text (neither in text nor parenthetical); however, the sentence uses identifiable themes from the source text and direct mention of the source text is used elsewhere in the essay. As a general rule, sentences that refer to ideas from a source text should include some direct mention of the source, and this lack of citation can amount to plagiarism: either deliberate or unintentional; however, after reviewing a large number of sentences in this analysis, there appear to be a number of exceptions to this general rule (examples of arguably reasonable indirect paraphrases are discussed more in the next paragraph). To explore the functional nature of this category, the researcher analyzed indirect paraphrases in context: first for L1 and then for L2 essays. Due to space limitations, this is the only type of source integration from Table 3 described in this way.

One pattern that emerged in L1 indirect paraphrases was a tendency to use what were identified as indirect paraphrases as a form of topic sentence or concluding remark. While the topic being introduced by the L1 student was clearly informed by the source text, no direct mention was made of the source, perhaps because the statement felt rather general in nature.

(1) *Another issue that can occur with international students is making new friends.* (Text 26)
(2) *Students coming to the United States for college face many challenges and obsatcles, such as making friends and adjusting to another culture*. (Text 40)
(3) *International students face an immense struggle when studying abroad. Preparation is imperative when doing so and they should remember to go outside of their comfort zones, learn of their surroundings, and join clubs and social activities to adjust to their new living arrangements.* (Text 46 - conclusion)

These statements serve as either a pre-cursor to more detailed information presented in subsequent sentences (Examples 1 & 2), or as an inductive statement in Example 3, serving as a summation of details previously discussed.

In terms of the L2 texts, indirect paraphrases were often involved in strings of sentences that included both directly attributed sentences and indirect paraphrases.

(4) *Two main issues that students coming to the United States have to deal with are cultural shock and the difficulty of making friends. (Tuna Asino). Making friends is becoming a major problem for international students.*[5] (Text 31)
(5) *This essay will show you main problems as well as solutions that help a lot of international students like us according to " nternational students encounter culture shock in a college aboard" and "how to survive as an international student". Firstly, the main problem is probably culture shock in college aboard.* (Text 138)

It should be noted that paraphrases with either in-text and parenthetical citations were also higher for L2 students in comparison to their L1 classmates (L2 = 3.19, L1 = 1.66). This impressionistic, qualitative analysis confirmed that L2 writers in this sample may have been more effective at citing sources *and* more effective at stringing references to source texts together in sentences that included at least one reference to the source text within the larger string of sentences. Yet, even with these relative strengths in referencing skills, L2 writers achieved significantly lower holistic quality scores, suggesting that L2 students' reading-to-write challenges may have more to do with other issues, perhaps including which ideational units were used and most importantly how they were incorporated into student texts.

In summary, differences in patterns of use for the *paraphrase indirectly referenced (pir)* category provide more context for the quantitative findings. For L1 writers, a rhetorical move that seemed to be occurring frequently in connection with this

[5] Sentences coded as *paraphrase indirectly referenced (pir)* are underlined.

code was to advance general summative statements. For L2 writers, these unreferenced sentences were often part of source-based sentence strings, with references present in adjacent sentences.

4.2. Research question 2: What patterns emerge in terms of which ideational units from the source texts are used by FYC L1 and L2 students on a 40-min, problem-solution, source-based writing task?

Table 4 includes normed frequency counts for the ideational unit drawn from the source texts (i.e., two source texts per prompt) and used in the students' writing. One general pattern that held true for both L1 and L2 source use was that both groups of writers were more likely to use ideational units from the first source than from the second source text in the pair (first source use = 36.21, second source use = 21.34).

When comparing L1 and L2 ideational unit use, different patterns emerged for different prompts. For the first prompt about international study abroad, the L1 writing used more ideational units in total when compared to L2 writing (L1 = 16.38, L2 = 12.37), and showed much stronger tendencies to use certain ideational units like *culture shock* (discussed in paragraphs 3 & 4 of Prompt 1, Source 1) and *post-arrival strategies* (discussed in paragraphs 3 & 4 of Prompt 1, Source 2) when compared to L2 writers. For the second pair of essays about global warming, however, the L2 students used more total ideational units than the L1 students (L1 = 12.08, L2 = 16.71), with the first source text (i.e., Prompt 2, Source 1) accounting for most of the observed differences in frequency. Also in contrast to the ideational unit use for the first prompt, the second prompt did not as strongly elicit patterns of use that favored certain ideational units over others. That is, the standard deviation for ideational unit use of both L1 and L2 essays written on the second prompt was considerably lower (SD = .84) compared to the standard deviation for ideational units used by both sets of writers on the first prompt (SD = 1.63).

This analysis of ideational units also provided an opportunity to investigate synthesis writing. By identifying instances in which ideational units from both sources for a prompt were assigned to the same sentence (e.g., *culture shock* from Prompt 1, Source 1, and *post-arrival strategies* from Prompt 1, Source 2), the researcher could determine the amount of synthesis writing being performed. Only 2% of coded ideational units involved this type of synthesis writing, with identical normed frequencies for L1 and L2 writers (i.e., .66 instances of synthesis per 1,000 words).

The category *post-arrival strategies for reducing culture shock* (pas) was chosen for a follow-up qualitative analysis because it represented (a) the highest combined L1 and L2 frequency for an ideational unit in Prompt 1, Source 2, and (b) a clear group difference (L1 = 3.87, L2 = 1.56). Upon further analysis of this general topic, it became clear that three sub-themes were folded into this one category: (1) keeping in touch with your home country, (2) taking a tour, and (3) being courageous, meeting new people, making new friends.

For L1 students, this functional analysis of ideational units coded as *post-arrival strategies for reducing culture shock* (pas) indicated elaborative development with a disproportionate emphasis on just one of the three sub-themes: making of new friends. In addition to incorporating source themes, L1 writers in this sample appeared more prone to also include themes and concepts from outside of the source information.

(6) *In America, the best way to make in friends is to look around, find things that interest you, join clubs, and meet new people who are interested in said things* (Text 133)
(7) *After arriving it is good for students to get out and explore their new environment. Take some time away from the school work and go make new friends and be active in social activities.* (Text 10)
(8) *You might even be interested, in a safe manner, connecting with people who will be in your surroundings via internet.* (Text 102)

As shown in the above examples, L1 students were using the source texts for inspiration, but were extending the thematic unit of making new friends in ways not directly mentioned in the text (e.g., connecting via the internet).

Table 4
Normalized frequency counts (per 1,000 words) for ideational units.

Topics & Code	Source	L1 Texts	L2 Texts
Culture shock (cs)	Prompt 1, Source 1	5.51	4.01
Difficulty making friends (dmf)	Prompt 1, Source 1	3.74	3.60
Classroom issues (ci)	Prompt 1, Source 1	1.02	0.74
Pre-departure strategies for reducing culture shock (pds)	Prompt 1, Source 2	1.19	1.56
Post-arrival strategies for reducing culture shock (pas)	Prompt 1, Source 2	3.87	1.56
Holistically in shape (hs)	Prompt 1, Source 2	1.04	0.90
Heating related (hr)	Prompt 2, Source 1	1.54	3.52
Weather effects (we)	Prompt 2, Source 1	1.83	2.70
Geological effects (ge)	Prompt 2, Source 1	1.33	2.87
Human health effects (hhe)	Prompt 2, Source 1	1.48	2.29
Renewable energy sources (res)	Prompt 2, Source 2	0.98	1.39
Recycling issues (ri)	Prompt 2, Source 2	1.46	1.15
Home conservation (hc)	Prompt 2, Source 2	0.62	0.25
Vehicle related (vr)	Prompt 2, Source 2	1.06	0.66
Reducing wasteful usage (rwu)	Prompt 2, Source 2	1.77	1.88

By contrast, L2 students approached this more general category of *post-arrival strategies for reducing culture shock* (pas) focusing disproportionately on a different sub-theme: keeping in touch with their home culture. Additionally, they seemed to provide considerably less elaboration, with shorter, listing of source text material. For example:

(9) *Joining club and organizations, and keeping tough with own family* (Text 130)
(10) *Introduce yourself and at the same time explore your environment …* (Text 134)
(11) *Moreover, international students should explore a new environment to reduce the pressure and also keep in touch with home culture.* (Text 141)

In all of the examples above, the connection with the source text is clear, yet the depth and elaboration of the themes presented are limited. A key difference between the L1 and L2 examples provided above is the degree to which information from the source text is *transformed* and expanded in support of the L1 writers' efforts, versus the degree to which the L2 writers are simply "knowledge telling," or shifting a thematic unit from source text to writer text with little change, growth, or repurposing (Bereiter & Scardamalia, 1987).

Shifting to the second prompt, the *heating-related* (hr) ideational unit was chosen for further analysis because it (a) had the highest combined L1 and L2 frequency for an ideational unit in the second prompt, and (b) had the largest between group frequency differences in the second prompt (L2 = 3.52, L1 = 1.54). This theme involved primarily facts and statistics about past heating trends and future global warming projections. In a prompt dedicated to problems and solutions in relation to global warming, *heating-related* issues were important to the text, in establishing evidence of warming as well predicting potential future problems.

For the L1 writers, it appears that the category of *heating-related* material from the source texts was explored as a general topic, which was then used as an entry point for other, related topics. For example,

(12) *Major issues we face today are heavy pollution in our cities, warmer temperatures, high water levels, outrageous weather patterns, etc. not only are we hurting our planet we are damaging ourselves.* (Text 106)
(13) *The global temperature has risen not only in the poles but everywhere else as well. Most notably Europe know to be a more cooler temperature area has been struck by massive heat waves that have altered Europe's agriculture industry and affected its tourism as well.* (Text 112)
(14) *Scientists say that the earth could warm by an additional 7.2 degrees Fahrenheit during the 21st century that if we fail to reduce emissions from burning fossil fuels, such as coal and oil, the rise in average temperature will affect the climate patterns of all living things* (Text 157)

As shown in Examples 12–14, warmer temperatures are one piece of multi-faceted sentences. In Example 12, the list of major issues concisely links warming with other issues such as pollution, water levels and weather patterns. Example 13 links warming with human health effects and agricultural production, while the final Example (14) makes connections to fossil fuels and energy issues.

By contrast, L2 writers used the issue of warming more than twice as frequently (L2 = 3.56, L1 = 1.54); however, the linking of ideational units to related themes was more limited. Once again, this underdevelopment of the L2 texts can be seen as a lack of elaboration. While a clear cause/effect relationship was more often present with these L2 writers, the relationship was rather simple. For example

(15) *In recent years, the weather is becoming warmer and warmer.* (Text 207)
(16) *Basic on global warming, the temperature becomes much higher than before.* (Text 81)
(17) *The first way is heat waves which can cause heat-related deaths* (Text 200)
(18) *The sea level and the temperature had risen these years, it is caused by the global warming.* (Text 207)
(19) *Because of the increasing of the temperature, sea level is raising.* (Text 11)

For Examples 15 and 16, the statements contain only information related to warming. For the remainder of the examples, a two-piece cause and effect structure is effective, but simple in terms of language use.

The functional analyses of the ideational units of *post-arrival strategies (pas)* and *heating related (hr)* both revealed issues of elaboration. For L1 writers, source material was often used to connect with related themes and to link with overlapping topics. For L2 writers, there was clear reporting of source-based material, especially when that material was concrete and statistical in nature, but little expansion and elaboration to related themes and concepts.

5. Discussion

The current study investigates L1 and L2 FYC student writing and the integration of source texts. Using the in-class writing samples, this exploratory analysis investigated two primary components for using source material: the types of source integration (e.g., *how* sources were integrated), and ideational units (e.g., *what* information was drawn from source texts). The analysis conducted in the current study indicated that L2 writers produced higher amounts of attributed paraphrases (both

in-text and parenthetical), as well as higher amounts of indirectly attributed paraphrases (meaning the source is attributed elsewhere in the student text) relative to the L1 writers. Additionally, the L2 writing included less paraphrasing with no source attribution relative to L1 writing. Both student groups used more source material from the first source, supporting findings from previous research (Gradwohl Nash et al., 1993).

The holistic quality scoring in the current study both provides context for the remaining analyses and adds points of comparison with previous studies. In the current study, the L1 writing achieved significantly higher scores than the L2 writing, with a rather large effect size. Previous L1 and L2 writing research demonstrates that student writing that achieves different holistic quality scores is likely to exhibit different patterns of source use (Campbell, 1990; Cumming et al., 2006; Gebril & Plakans, 2009, 2013); therefore, comparisons between these student groups and across studies should take into account this finding. In terms of previous research, Campbell (1990), in a study of 30 university students, also found that L1 writers achieved significantly higher holistic quality scores than either the high or low proficiency L2 writers. Counter-evidence may be found in Asención Delaney (2008), who, using an analytic rubric, indicated no significant differences between L1 and advanced L2 students, though an argument could be made that the students in the current study more closely resembled the intermediate L2 students from her study (paper-based TOEFL scores of 450–543), a group which did produce significantly lower writing quality scores than the writing of L1 students.

Mixed-methods findings from the current study provide interesting comparisons with previous studies in terms of students' types of source integration. Support for the finding that L2 students are more responsible with citation practices can be found in Campbell (1990), who indicated that L2 writers were more than twice as likely to include citations for quotes and summaries than L1 writers. The current findings stand in contrast, however, to Shi (2004), who found that L2 writers (with L1 Chinese) were less likely to cite sources than their L1 classmates, with this pattern holding true for both longer strings of verbatim copying and short strings (less than six words). Also, while Lee et al. (2018) investigated only cited source use and identified 8.48 occurrences of some form of paraphrase, summary, or synthesis per 1,000 words for L2 writers, the current study also included uncited and indirectly cited source use and found considerably higher frequencies, with 23.59 occurrences per 1,000 words for L2 writers. Studies focusing exclusively on student citation practices provide valuable contributions to the study of source-based writing (Howard et al., 2010; Jamieson, 2013; Lee et al., 2018), but interestingly, if the current study had excluded non-cited source use, it would not have accounted for approximately 79% of all source uses. As shown in the functional analysis, L1 writers used indirect paraphrasing as a form of transition, and L2 writers paraphrased indirectly within strings of sentences where the citation is provided elsewhere. Both of these patterns indicate a potential gray area in the necessity of providing citation material (Shi, 2008). One implication from the current study involves this gray area of indirectly and uncited source support. Educators, students, and researchers would do well to more clearly understand (a) how common indirectly and uncited source use is, (b) when source citation is required, and (c) when using source-ideas without including a citation is widely-accepted.

Another important finding from the current study was the lack of synthesis in both L1 and L2 writing. Because students wrote from two sources and were instructed to "combine information from both texts" in their writing, one might expect the student to have produced a reasonable amount of synthesis writing. Yet, in sentences with ideational units from one source, only 2% of those sentences included ideational units from the other source as well. As a comparison, taking an average across eight academic disciplines, Hyland (2000) found that 27% of expert writer citations included multiple sources. One important difference between the current study's 2% synthesis and Hyland's 27% synthesis writing was the 40-min time constraint in the current study, versus the unlimited revision time available to the experts. Nevertheless, taking into account the time provided, this reluctance and/or difficulty for students to synthesize source material across texts is consistent with previous L2 (Plakans, 2009a; Zhao & Hirvela, 2015) and L1 research (Sole et al., 2013). Facilitating greater elaboration in L2 synthesis writing presupposes the skills of summary writing (Hirvela, 2016), and also requires what Zhao and Hirvela (2015) call rhetorical reading, when "readers construct meanings of texts by considering the context, textual cues, the author's purpose, and audience expectations" (p. 231) as well as task representation, or "interacting with the text to construct his/her own meaning" (p. 229). While providing students more time for the writing task may have improved the synthesis writing, neither the L1 nor the L2 students excelled in synthesis writing in the current study and a deliberately designed curriculum emphasizing synthesis writing would be an effective way to build these synthesis skills in FYC writers. Perhaps FYC courses could place a greater emphasis on "hybrid literacies" in which reading and writing work together in overlapping and recursive ways (Spivey, 1990, p. 259).

In terms of the ideational units drawn from the source texts, for both of the categories analyzed more closely (e.g., *post-arrival strategies* and *heating-related*) it appears that the L2 writing involved more "knowledge telling" (Bereiter & Scardamalia, 1987), or simple reporting of source material with support from the writer's background knowledge of the topic and understanding of the expectations of an informative essay. L1 writers on the other hand were more likely to integrate the source material into sentences that extended beyond one chunk of material from the source text. This L2 writing, marked by "knowledge telling" and lacking in elaboration, is consistent with much previous research (Gebril & Plakans, 2009; Li & Casanave, 2012; Wette, 2017a). Wette (2017a) frames this as an issue of using sources for a limited range of rhetorical purposes. She suggests that, while the L1 writers are using the sources to enhance their personal perspectives on the prompt, L2 writers are using the source material as a stand-alone "proposition, theory, fact, research finding, or item of information" (p. 53). It may be the case that, despite the strengths of L2 writing compared with L1 writing regarding the use of citations, this lack of elaboration was an important contributor to the significantly lower holistic quality scores for the L2 writing compared to the L1 writing.

This exploratory analysis identified a series of writing patterns uncovered under controlled writing conditions and drawn from just two prompts and four short source readings. Writing instructors, on the other hand, often facilitate student source-based writing from considerably *more* source texts and under considerably *less* controlled writing conditions. Educators are unlikely to have the time to analyze patterns of referencing, quotation, paraphrases, and synthesis such as those undertaken in the current study. Yet, teachers could certainly use empirical research studies such as this one to incorporate a small set of source-based writing units into their FYC courses (Zhang, 2013). Instructors could focus as a class on a small number of pedagogical issues generated from the current study such as simple versus elaborate source use, citation practices in strings of source-based material, and synthesis creation.

6. Limitations

While the current study was carefully designed, it was not without its limitations. First, achieving a higher inter-rater reliability for the holistic quality scoring would have been preferable. Second, while the number of L1 participants was large ($n = 112$), a larger L2 group ($n = 37$) would have been optimal. Especially with student writing split between the two writing prompts, a more robust L2 population could have strengthened quantitative comparisons being made. Also, in terms of the prompts, while Zhang (2013) makes the case that the prompts are comparable, it is noteworthy that L1 writers used more ideational units in the first prompt compared to the L2 writers while the L2 writing included more ideational units from the second prompt compared to the L1 writing. A more fine-grained analysis of source text differences was beyond the scope of this study, but might provide additional and valuable context.

The discourse analysis coding scheme was developed for the current study, and future efforts to refine these coding procedures are likely to strengthen the reliability and validity of these analyses. The types of source integration coding scheme involved categories in and around the gray areas of plagiarism, and persuasive arguments can be made by those who adopt different perspectives on whether or not writers made good-faith efforts at citing appropriately (cf., the Disciplinary Dialogues section of the *Journal of Second Language Writing*, Pecorari, 2015). Specifically, the category of *paraphrase indirectly referenced* can be viewed by different instructors as ranging from clear plagiarism at one end of the continuum, to appropriate and productive at the other (Pecorari & Shaw, 2012). Additionally the category of "plagiarism possible" involves a rather imprecise definition, and the reader can determine how much (and what type of) value to ascribe to this category.

Finally, though not exactly a limitation, readers should be reminded that the study of in-class writing, completed in 40-min, involves a specific set of situational characteristics that have benefits and drawbacks relative to the study of take-home writing. While take-home writing introduces problematic moderating variables for research purposes (e.g., time on task, distractions in the environment, assistance), it is clearly the more frequently occurring type of writing in FYC classes.

7. Conclusion

The current study represents an exploratory, descriptive overview of the integration of source material among L1 and L2 student writers on an in-class, source-based writing task. By analyzing students' types of source integration and their selection of ideational units, the current study explored similarities and differences between these student groups and qualitatively investigated patterns that emerged. It should be recognized that the writing quality of L2 students when compared to L1 students in the current study was significantly lower. That said, this study does not just focus on deficiencies in L2 writing. Despite lower writing proficiencies, L2 writers demonstrated strengths in citation practices (e.g., directly and indirectly attributed paraphrases and quotations with reference), stringing together source-based material, and a willingness to incorporate more source-based information from the second prompt than the L1 writers.

This study also represents an important contribution to our understanding of source-based writing for both pedagogy and research. First, as discussed above, source-based writing is required across much of the university; therefore, building FYC classes (for both L1 and L2 students) around writing from sources is a way to better prepare students for the writing they will need to do in their future studies (Doolan & Fitzsimmons-Doolan, 2016). Yet, for teachers to build an FYC curriculum attending to source-based writing issues, more studies are needed such as this one that identify the skills associated with the integration of source material at the beginning of the FYC semester. From a research perspective, the current study helps provide a greater understanding of students' source-based writing skills coming *into* FYC classes, which can help researchers reconcile performance at this stage of development with the wealth of research on source-based writing as students advance through later stages of their college careers. This combined knowledge advances greater cross-sectional understanding of source-based writing development.

For future research and pedagogy, the inclusion of multiple source texts in writing tasks introduces complex variables that have the potential to impact writing performance in ways we do not yet fully understand. As research informs pedagogy, it may be worthwhile to further classify the quality of source text use (along with the presence of source text use) in order to help build a preliminary understanding of how source text use develops over time (Wette, 2017b). For those building pedagogy around source-based writing, interventions can be developed that do not just focus on specific areas of difficulty in L2 writing from sources, but also take into account empirically observed strengths that L2 writers bring to the classroom. As source-based writing research continues to evolve, scholars can establish a more empirical, descriptive understanding of how students reference sources and what material students are prone to use from sources. Such developments will surely benefit the classroom.

Acknowledgments

I would like to thank Dr. Shannon Fitzsimmons-Doolan for extensive feedback on this article. Thank you also to Mary Hood, Chimene Jimena, Clarissa Reeves, and Tasha Aschmutat for assistance with data analysis. Finally, thank you to the editor and to the anonymous reviewers for their attention to detail and suggestions that certainly improved this article.

Appendix A

Source-Based Writing Rubric (From Zhang, 2013, Adapted from Plakans, 2009b).

Score	Description
5	A response at this level: • Successfully addresses the assignment through the use of a clear problem-solution organizational pattern • Successfully presents all important information from source texts in relation to prompt • Is well organized with well-developed content • Occasional language errors that are present do not result in inaccurate or imprecise presentation of content or connections
4	A response at this level: • Adequately addresses the assignment through the use of a relatively clear problem-solution organizational pattern • Is generally good in coherently and accurately presenting relevant information from source texts, although the response may have information not (correctly) referenced • Has clear organization and logical development • Contains more frequent or noticeable minor language errors that do not result in anything more than an occasional lapse of clarity
3	A response at this level: • Largely addresses the assignment; problem-solution organizational pattern is present but may need reader effort to identify • Presents most important information from source texts, but may be lack of one or two problems/solutions, or problems and solutions may be mismatched • Occasionally lacks cohesion but has a basic organizational structure and development • Includes many usage and grammar errors that may result in noticeably vague expressions or obscured meanings
2	A response at this level: • Partially addresses the assignment; only parts of problems and/or solutions are present • Contains some relevant information from the readings, but is marked by significant omission or inaccuracy of important ideas from the readings or largely copied texts • Lacks logical organizational coherence and development • Contains language errors or expressions that largely obscure connections or meaning at key junctures, or that would likely obscure understanding of key ideas
1	A response at this level: • Addresses the assignment to a very limited degree; problem-solution pattern is not evident • Provides little or no meaningful or relevant coherent content from the readings and does not follow an organization pattern or develop content • Most language in the writing is copied or includes language that is so low it is difficult to derive meaning
0	A response at this level: • Either merely copies sentences from the reading, rejects the topic, not connected to the topic, is written in a foreign language, or is blank

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.esp.2021.01.003.

References

Asención Delaney, Y. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes, 7*, 140-150.

Basturkmen, H. (2017). Developing writing courses for specific academic purposes. In J. Flowerdew, & T. Costley (Eds.), *Discipline-specific writing* (pp. 31-45). New York, NY: Routledge.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Erlbaum.

Campbell, C. (1990). Writing with others' words: Using background reading text in academic composition. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 211-230). New York, NY: Cambridge.

Carson, J. G. (2001). A task analysis of reading and writing in academic contexts. In D. Belcher, & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 48-83). Ann Arbor, MI: University of Michigan Press.

Cumming, A., Kantor, R., Baba, K., Eouanzoui, K., Erdosy, K., & James, M. (2006). *Analysis of discourse features and verification of scoring levels for independent and integrated prototype written tasks for the new TOEFL* (Monograph Series, No. 30). Princeton, NJ: Educational Testing Service.

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes, 23*, 47-58.

Currie, P. (1998). Staying out of trouble: Apparent plagiarism and academic survival. *Journal of Second Language Writing, 7*, 1-18.

Dollahite, N., & Haun, J. (2012). *Sourcework: Academic writing from sources* (2nd ed.). Boston, MA: Cengage.

Doolan, S. M. (2017). Comparing patterns of error in Generation 1.5, L1, and L2 first-year composition writing. *Journal of Second Language Writing, 35*, 1-17. https://doi.org/10.1016/j.jslw.2016.11.002.

Doolan, S. M., & Fitzsimmons-Doolan, S. (2016). Facilitating interpretation of source texts for L2 writers. *TESOL Journal, 7*, 716-745. https://doi.org/10.1002/tesj.239.
Gebril, A., & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 7*, 47-84.
Gebril, A., & Plakans, L. (2013). Toward a transparent construct of reading-to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly, 10*, 9-27.
Grabe, W., & Zhang, C. (2013). Reading and writing together: A critical component of English for academic purposes teaching and learning. *TESOL Journal, 4*, 9-24. https://doi.org/10.1002/tesj.65.
Gradwohl Nash, J., Schumacher, G. M., & Carlson, B. W. (1993). Writing from sources: A structure-mapping model. *Journal of Educational Psychology, 85*, 159-170.
Hirvela, A. (2016). *Connecting reading & writing in second language writing instruction* (2nd ed.). Ann Arbor, MI: University of Michigan Press.
Howard, R. M., Serviss, T., & Rodrigue, T. K. (2010). Writing from sources, writing from sentences. *Writing & Pedagogy, 2*, 177-192. https://doi.org/10.1558/wap.v2i2.177.
Hyland, K. (2000). Academic attribution: Interaction through citation. In K. Hyland (Ed.), *Disciplinary discourses: Social interactions in academic writing* (pp. 20-40). Harlow: Pearson.
James, M. A. (2018). Teaching for transfer of second language learning. *Language Teaching, 51*, 330-348.
Jamieson, S. (2013). Reading and engaging sources: What students' use of sources reveals about advanced writing skills. *Across the Disciplines, 10*. Retrieved February 1, 2018, from http://wac.colostate.edu/atd/reading/jamieson.cfm.
Johns, A. M., & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics, 11*, 253-271.
Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing, 25*, 4-22. https://doi.org/10.1016/j.jslw.2014.05.005.
Kim, S. (2001). Characteristics of EFL readers' summary writing: A study with Korean university students. *Foreign Language Annals, 34*, 569-581. https://doi.org/10.1111/j.1944-9720.2001.tb02104.x.
Lee, J. J., Hitchcock, C., & Casal, J. E. (2018). Citation practices of L2 university students in first-year writing: Form, function and stance. *Journal of English for Academic Purposes, 33*, 1-11.
Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31*, 39-69.
Li, Y., & Casanave, C. (2012). Two first-year students' strategires for writing from sources: Patchwriting or plagiarism? *Journal of Second Language Writing, 21*, 165-180. https://doi.org/10.1016/j.jslw.2012.03.002.
Liu, G., Lin, V., Kou, X., & Wang, H. (2016). Best practices in L2 English source use pedagogy: A thematic review and synthesis of empirical studies. *Educational Research Review, 19*, 36-57.
McGinley, W. (1992). The role of reading and writing while composing from sources. *Reading Research Quarterly, 27*, 226-248.
Pecorari, D. (2015). Disciplinary dialogues. *Journal of Second Language Writing, 30*, 94-110.
Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing, 21*, 149-164.
Plakans, L. (2009a). Discourse synthesis in integrated second language writing assessment. *Language Testing, 26*, 561-587. https://doi.org/10.1177/0265532209340192.
Plakans, L. (2009b). The role of reading strategies in integrated L2 writing tasks. *Journal of English for Academic Purposes, 8*, 252-266. https://doi.org/10.1016/j.jeap.2009.05.001.
Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing, 17*, 18-34. https://doi.org/10.1016/j.asw.2011.09.002.
Polio, C., & Friedman, D. (2016). *Understanding, evaluating, and conducting second language writing research*. New York NY: Routledge.
Risemberg, R. (1996). Reading to write: Self-regulated learning strategies when writing essays from sources. *Reading Research & Instruction, 35*, 365-383.
Rosenfeld, M., Leung, S., & Oltman, P. (2001). *The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels* (TESOL Monograph Series MS-21). Princeton, NJ: Educational Testing Service.
Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication, 21*, 171-200. https://doi.org/10.1177/0741088303262846.
Shi, L. (2008). Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics, 31*, 1-24.
Sole, I., Miras, M., Castells, N., Espino, S., & Minguela, M. (2013). Integrating information: An analysis of the processes involved and the products generated in a written synthesis task. *Written Communication, 30*, 63-90.
Spivey, N. N. (1990). Transforming texts: Constructive processes in reading and writing. *Written Communication, 7*, 256-287.
Spivey, N. N., & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly, 24*, 7-26.
Storch, N. (2012). Incorporation of source material: The effect of instruction. *TESOL in Context, 22*, 38-55.
Stotsky, S. (1983). Research on reading/writing relationships: A synthesis and suggested directions. *Reading and Writing, 60*, 627-642.
Sun, Y. (2012). Does text readability matter? A study of paraphrasing and plagiarism in English as a foreign Language Writing context. *The Asian-Pacific Education Researcher, 21*, 296-306.
Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing, 9*, 27-55. https://doi.org/10.1016/j.asw.2004.01.002.
Wette, R. (2017a). Source text use by undergraduate post-novice L2 writers in disciplinary assignments: Progress and ongoing challenges. *Journal of Second Language Writing, 37*, 46-58.
Wette, R. (2017b). L2 undergraduate students learning to write using sources: A trajectory of skill development. In J. Bitchener, N. Storch, & R. Wette (Eds.), *Teaching writing for academic purposes to multilingual students* (pp. 99-114). New York, NY: Routledge.
Zhang, C. (2013). Effect of instruction on ESL students' synthesis writing. *Journal of Second Language Writing, 22*, 51-67. https://doi.org/10.1016/j.jslw.2012.12.001.
Zhao, R., & Hirvela, A. (2015). Undergraduate ESL students' engagement in academic reading and writing in learning to write a synthesis paper. *Reading in a Foregin Language, 27*, 219-241.

Dr. Stephen Doolan is an Associate Professor of Applied Linguistics at Texas A&M University – Corpus Christi. His research interests include L2 writing, resident L2 writing, and source-based writing. Dr. Doolan has recently published in the *Journal of Second Language Writing,* the *TESOL Journal,* and *Written Communication*.