



Castledown

 OPEN ACCESS

Language Education & Assessment

ISSN 2209-3591

<https://www.castledown.com/journals/lea/>

Language Education & Assessment, 4(3), 81–104 (2021)
<https://doi.org/10.29140/lea.v4n3.507>

Exploring EAP students' perceptions of integrated writing assessment



PAKIZE ULUDAG

Concordia University, Canada
pakize.uludag@concordia.ca

Abstract

Despite the breadth of integrated writing assessment research, few studies have examined student perceptions of classroom-based integrated writing tasks or the instructional value of analytic rubrics. Adopting a case study methodology, this exploratory research investigated three L2 learners' conceptualizations of integrated writing assessment and their use of an analytic rubric for self-assessment in an EAP writing course. Data sources included integrated writing samples that were evaluated by the students and their instructor, a writing self-efficacy questionnaire, individual retrospective interviews, and course materials such as syllabi, and task instructions. Qualitative analysis revealed themes related to three aspects of classroom-based integrated writing assessment: task requirements, task conditions, and instructor feedback. The themes were discussed in terms of students' test taking strategies and the use of available support systems in EAP contexts. In addition, findings indicated an overlap between students' self-assessment and instructor evaluation of their integrated essays, suggesting that students could use the evaluation criteria effectively. Implications for teaching and assessing integrated writing in EAP contexts are discussed.

Keywords: integrated writing assessment, EAP, student perceptions, case study

Introduction

Integrated writing tasks have become a standard component of large-scale English proficiency tests, such as the Test of English as a Foreign Language Internet-Based Test (TOEFL iBT), and the Canadian Academic English Language (CAEL) test to ensure the validity and authenticity of writing assessment. These tasks typically require that writers incorporate information from source materials into their own writing (Cumming et al., 2006). To write an essay using information from sources, students must understand the arguments presented in the sources, and revise source-text language without distorting the original meaning (Plakans, 2008, 2009). Reflecting the strong emphasis on academic writing in university courses, English for Academic Purposes (EAP) programs have also

Copyright: © 2021 Pakize Uludag. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within this paper.

incorporated integrated writing in their curriculum and assessment (Plakans, 2009, 2010; Plakans & Gebril, 2012) to help students transition to disciplinary writing. The use of integrated writing tasks for EAP pedagogy and standardized writing tests has not only improved the measures of writing ability, but also contributed to students' engagement with academic tasks that are integral to disciplinary contexts in higher education (Cumming *et al.*, 2006).

During the last few decades, second language (L2) writing researchers have explored the link between observed scores and integrated writing quality to address the validity issues associated with the use of integrated writing tasks. Few studies have focused on rater perceptions to examine the scoring validity and standardization of rater procedures (Cumming *et al.*, 2002; Gebril & Plakans, 2014). In addition, conducting textual analysis, researchers have demonstrated that both reading and writing abilities are pertinent to integrated writing test performance (Leki & Carson, 1997; Watanabe, 2001). Although these studies have shed light on the factors that impact the assessment of integrated tasks, it remains unexplored how students conceptualize classroom-based integrated writing assessment and whether they can use integrated writing assessment criteria for self-assessment. Therefore, to establish clearer assessment standards and improve integrated writing instruction in an EAP context, this case study, as part of a larger study about the design and validation of an integrated writing rubric, explored students' insights into classroom-based integrated writing assessment and their use of an analytic rubric for self-assessment.

Evidence of test validity from the perspectives of stakeholders

One of the concerns in the field of language testing is the development of tasks that represent the characteristics of the "target language use (TLU) domain" (Bachman & Palmer, 1996), which refers to the "situation or context in which the test taker will be using the language outside of the test itself" (p. 18). Researchers working within the assessment for learning perspective have increasingly recognized the importance of incorporating stakeholder perspectives into the development of tasks that reflect real-life domains (e.g., Rosenfeld *et al.* 2001; Malone & Montee, 2014). Studies focusing on the use of formative and ongoing assessments have reported that incorporating student feedback into revision of assessment tasks and test rubrics positively impacts teaching. In addition, articulation of students' thoughts and experiences with assessment has been shown to enhance student motivation and promote self-regulated learning (Benson, 2007; Butler, 2016; Rea-Dickins, 2006).

Although studies from the perspectives of EAP students have been limited, researchers have explored how teachers and assessment professionals understand assessment constructs as compared to student understanding. For example, Sato and Ikeda (2015) explored Japanese and Korean college students' interpretations of the skills targeted in a high-stakes college entrance examination that used a multiple-choice format for assessing receptive and productive language skills. Students did not have a clear understanding of test items that measured the ability to read between the lines and writing ability. Similar findings have been reported for the use of alternative assessment methods in different EFL contexts. For example, Vlanti (2012) revealed discrepancies between the interpretations of the Greek junior high-school students and their instructors in terms of the use of self- and peer-assessment for grading in language classrooms. Thus, researchers have emphasized student involvement in assessment decisions to obtain positive washback on language teaching and learning.

More recently, studies have incorporated student perceptions into language assessment literacy (LAL) discussions (Butler *et al.*, 2021). This strand of research has demonstrated that students are major stakeholders (Erickson & Gustafsson, 2005) who can reflect on and influence the assessment processes (Butler, 2019). For example, Butler *et al.* (2021) investigated fourth- and sixth-grade

Chinese students' knowledge and understanding of assessment purposes, skills, and principles drawing on existing LAL models. Administering a mock English test and eliciting student perceptions through semi-structured interviews, they found that students had substantial prior knowledge and experience with assessment. In addition, despite being young, they were capable of articulating their wants and needs (e.g., communicative-based assessment) and identifying construct-irrelevant factors, such as anxiety, which might affect their test performance.

Situated within this line of research, a number of researchers have elicited student perceptions to confirm the validity of large-scale English language tests. For example, Winke et al. (2018) interviewed both L1 and L2 English speaking children (ages 7 – 9) to investigate the cognitive validity of the Young Learners Tests of English administered by Cambridge Michigan Language Assessments. They found that incorrect responses by L1 children resulted from age-related cognitive constraints and lack of assessment literacy, which were both considered to be construct-irrelevant variance. Similarly, Cheng and DeLuca (2011) incorporated student perspectives into validation of four different high-stakes language tests used for entrance and certification at an English-medium university in Asia. The students' testing experiences concerned both construct representation (e.g., testing consequences) and construct-irrelevant variance (e.g., test format and administration). Both studies have confirmed that insights from test-takers contribute to test validity and help promote a more balanced and comprehensive understanding of LAL.

Focusing more specifically on writing tasks, researchers have compared the TOEFL iBT writing section and target language use (TLU) tasks (i.e., academic writing tasks in university courses) drawing on interview data from teachers and students. For example, Malone and Montee (2014) used stimulated recall interviews to elicit university students' perceptions of TOEFL iBT items, including the integrated writing task. They found that the majority of the students were able to identify the similarities and differences between TOEFL integrated writing tasks, which requires writing based on source material only, and the type of tasks they previously encountered in university courses. In a follow-up study, which focused specifically on TOEFL iBT writing tasks, Llosa and Malone (2017) obtained student and instructor perceptions, through which they revealed a convergence between the abilities assessed in TOEFL writing tasks and underlying TLU tasks, such as summary and synthesis writing. Taken together, assessment validation studies have confirmed the importance of obtaining student perspectives to enhance validity arguments and understand the outcomes of large-scale English language tests.

Integrated writing assessment in educational contexts

Despite the growing interest in classroom-based assessment, student perceptions of integrated writing test characteristics and scoring criteria have rarely been explored in previous research. Studies of L2 integrated writing have largely examined the individual and contextual factors that might play a role in students' writing processes. For example, Plakans (2009) used think-aloud protocols and interviews to examine L2 students' integrated writing processes on an ESL placement test. The results indicated that students' composing processes were influenced by their prior experiences and background knowledge of writing along with their topic familiarity. In a qualitative case study, Zhu (2005) investigated a graduate Chinese student's experiences with integrated writing tasks and found that the student's task representation was shaped by task purposes. In addition, the opinions expressed in the source texts facilitated the student's idea development and aided the direction of the responses. The function of source texts was also investigated by Plakans and Gebril (2013) using a mixed method design, which included interviews and questionnaires. The study found that students across different proficiency levels used source texts for similar purposes, such as developing personal opinions about the topic and for language support. Although not focused on

assessment constructs and uses, L2 students' perceptions and practices were also elicited in numerous case studies to explore source use variables, such as patchwriting and verbatim copying (e.g., Harwood & Petri, 2012; Hirvela & Du, 2013; Li & Casanave, 2004). Even though these studies have provided useful information related to students' textual borrowing and citation behavior while working on integrated writing tasks, they do not shed light on how students conceptualize integrated writing assessment or how they benefit from assessment tools, such as analytic rubrics, in classroom settings.

Research on evaluation criteria and self-assessment

Another important issue that merits attention from researchers concerns L2 students' attitudes and reactions to the use of integrated writing rubrics in EAP classrooms. Within the cycle of test development and validation, rubrics are reporting mechanisms that show students how their work is assessed and what skills they need to achieve specific success levels (Ferris & Hedgcock, 2014; Hamp-Lyons 2014). Describing language learning objectives and accomplishments in qualitative terms, evaluation criteria not only serve the purpose of transparency in assessment (Crusan, 2010, 2015; Hudson, 2005) but also help students become autonomous and responsible for their own learning by means of self-monitoring (Jonsson, 2014; Panadero & Jonsson, 2013).

Researchers exploring the instructional role of evaluation criteria in writing classes have suggested rubrics may constitute corrective feedback (Hyland, 2003) and help students improve their revision strategies in multi-drafted writing (Jonsson & Svingby, 2007). In addition, interaction-oriented assessment practices, such as self-assessment, have been shown to enhance students' intrinsic motivation and self-confidence, thereby facilitating self-regulated learning (Cheng & Fox, 2017; McMillan & Hearn, 2008; Yu & Jin, 2014; Wallace, 2018). However, apart from few rubric validation studies (e.g., Uludag & McDonough, 2022), research focusing on students' engagement with the evaluation criteria is scant in the area of integrated writing assessment. In fact, most discussions on integrated writing rubrics have concerned the reliability of the assessment criteria (Knoch *et al.*, 2007), with little attention to whether rubrics help students identify their strengths and weaknesses with source integration. Thus, it is important to explore to what extent instructor-oriented rubrics encourage self-monitoring and facilitate students' interpretation of test results (Alderson & Hamp-Lyons, 1996; Cheng, 2005). Establishing such evidence for the usefulness and practicality of evaluation criteria could promote positive washback and facilitate student participation in integrated writing assessment.

The Current Study

In summary, student perspectives have contributed to test validation and score interpretation in educational assessment contexts, such as large-scale achievement or proficiency tests. On the other hand, although research on students has yielded an increased recognition of the key role that students play as stakeholders, classroom-based assessment constructs and procedures have been rarely explored from students' perspectives. In particular, the research is scant on students' beliefs, experiences, and expectations related to classroom-based integrated writing assessment and evaluation criteria. Therefore, this exploratory case study focuses on students' experiences with and perceptions of integrated writing tasks in an EAP writing course, which required them to integrate source information with their personal opinions, and their use of an analytic rubric for self-assessment.

To gain a deeper understanding of the students' perceptions of integrated writing assessment, this study adopted an inductive case study approach. A case study design, commonly used in educational

research, involves an in-depth analysis of the cases in a real-life context and draws meaningful conclusions by taking an interpretive orientation (Simons, 2009; Yin, 2014). Placing an emphasis on the process (e.g., task design, teacher feedback) as well as the outcomes, the current study examined perceptions of three EAP students about integrated writing assessment by engaging multiple data sources in an organized and systematic way. The researcher, who had no supervisory or administrative role in the EAP program, acted as a participant-observer in a naturalistic EAP setting (Patton, 2002). The research questions are as follows:

1. What are students' perceptions of integrated writing assessment in an EAP course?
2. How well can EAP students apply an integrated writing rubric for self-assessment?

Methodology

Study context

The study was situated in an EAP program at an English-medium university in Canada, where two six-credit academic writing courses are offered to students. While the first course (Course 1) focuses on paragraph-level writing skills with a strong emphasis on vocabulary and grammatical development, the second course (Course 2) targets the analytical skills needed for integrated writing. L2 speakers who do not meet the English language proficiency test requirements for admission register for the EAP courses concurrent with their respective undergraduate degree programs (Table 1). Some students are exempted from Course 1 based on their performance on an in-house placement test. The current study took place in Course 2, which meets twice a week for 2.75 hours per meeting over a 13-week semester. The course objectives are to introduce students to source-based writing tasks, comprised of cause-and-effect, and argumentative essays, which require them to synthesize academic content, and develop and support their views on academic subjects.

Table 1 *English proficiency admissions requirements*

Test	Admission without EAP courses	Admission with EAP courses
TOEFL iBT	90 or higher with no component score under 20	75-89 with combined speaking and writing score of 34 or higher
IELTS	6.5 with no component scores under 6.5	Overall score of 6.5 with no component scores under 6.0

Participants

The study was announced to the students registered in the same section of the EAP writing course (Course 2) during the first week of the Winter 2021 semester. The course, originally designed for face-to-face instruction, was delivered in a synchronous mode using an online meeting tool (Zoom) due to the pandemic. Out of 24 students registered in the course, three students, whose names were replaced by pseudonyms (Liu, Qian, and Moza) agreed to participate in this case study. After reading and signing the consent forms, they completed a background questionnaire, which elicited specific information about their English learning background (Table 2).

Instructional Design

Pedagogical materials utilized in Course 2 were compiled in the course-pack by the EAP program coordinators and consisted of a) theme-based academic texts, b) authentic articles from newspapers

and news magazines, and c) vocabulary and grammar topics which led to unit-final integrated writing tasks. The theme-based academic texts accompanied by vocabulary exercises were selected based on their content. Authentic articles were intended to offer different perspectives or types of information that students could use when composing integrated essays on relevant topics. As for the writing skills, students received explicit instruction on how to paraphrase and summarize source information, in addition to learning citation skills using the APA publication manual (American Psychological Association, 2020).

Table 2 *Participants' background*

Name*	Liu	Qian	Moza
L1	Mandarin	Mandarin	Arabic
L2	English	English	French
Gender	Female	Male	Female
Age (years)	45	34	18
Length of residence in Canada	13 years	1.5 years	8 months
Length of English study in home country	12 years	10 years	13 years
Length of English study in Canada	2 years	1 year	8 months

Throughout the 13-week semester, students carried out several individual and group assignments, such as graded vocabulary and grammar quizzes, summary and paraphrasing exercises, and ungraded quizzes, which were planned as formative assessment. As for the EAP program's regularly-scheduled battery of summative tests, the students wrote a midterm exam, which elicited summary writing (approximately 300 words), and three classroom-based integrated writing tests (approximately 500 words each), which targeted two different genres (i.e., cause-and-effect and argumentative), and required them to integrate information from written sources and acknowledge the use of these sources through in-text citations.

Following the assessment protocol in the EAP program, two weeks prior to each exam, students were assigned a reading list with six to seven sources from the course-pack. They discussed these readings with the course instructor and prepared notes using a note-taking template. Students could refer to the note-taking sheet in the course-pack which is filled in with notes on a source reading from the same course-pack. This provided them with an example of how to transfer information from sources to the note-taking sheet. On the exam day, the students were given two integrated writing prompts and chose which one to write about. They were allowed to use one note-taking sheet per source as well as an English-only dictionary.

Data sources

To identify the students' perceptions of classroom-based integrated writing assessment, the researcher, as a participant-observer, was granted access to the online course page which included pedagogical materials, such as course syllabi, rubric, and task instructions. After writing the cause-and-effect integrated writing test (week 9), the students who volunteered for this study were asked to complete a writing self-efficacy questionnaire adapted from Abdel-Latif (2015), which contained 18 items (Appendix A). There were 8 items in the first section to assess students' judgement about their writing skills using a 5-point Likert scale (1 = strongly agree; 5 = strongly disagree). The second section with 10 items asked students to rate how confident they are when performing various tasks when writing using a 5-point Likert scale. Scores could range from 18 to

90, with high scores indicating high self-efficacy. Along with the self-efficacy questionnaire, the students answered some open-ended questions about the EAP writing course (see Appendix A) to obtain an understanding of their conceptualizations of the formative and summative assessment tasks prior to meeting with the researcher for an interview.

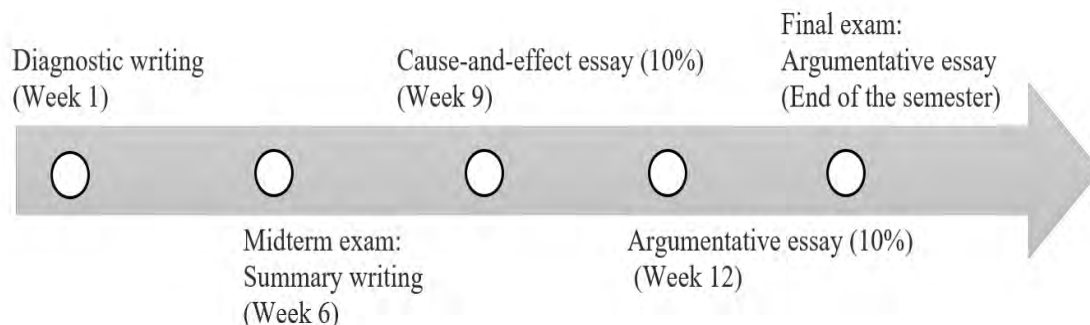


Figure 1 *Timeline for the summative assessment tasks*

After administering the online questionnaire, the researcher scheduled an individual data collection session with each student. During this meeting, a semi-structured interview was conducted as an initial step, which lasted approximately 60 minutes. The introspective open-ended questions elicited students' perceptions about the integrated writing assessment tasks, teacher feedback and their use of available support systems in the EAP program. In addition, information was collected about students' challenges, and strategies while writing from sources (e.g., generating ideas, selecting information from sources, incorporating citations, structuring the essay, and language-related reflections). To facilitate recall of their thought processes specific to source integration, the researcher referred to students' cause-and-effect essays, which they had recently completed as part of the summative assessment. Before the interview, the essays had been analyzed for the aspects of source use (i.e., accurate representation of source ideas, linguistic revision of source language, and source use purposes).

In the second part of the interview, students were asked to use an integrated writing rubric (Appendix B) to self-assess their cause-and-effect essay performance. The researcher introduced the rubric categories (i.e., content, organization, source use, and language use) and asked students to work independently to apply the rubric to their own writing, assigning a score for each category. Importantly, this rubric had been designed and validated as part of an ongoing research project in the EAP program (Uludag & McDonough, 2022), and it was different from the current EAP program rubric in terms of targeting source use as a separate category and including more detailed and task specific descriptors. After completing the self-assessment, the students showed their annotated rubrics and the essay to the researcher via the share screen function and discussed their scoring decisions. Notably, the interviews were conducted entirely in English and the students seemed comfortable while interacting with the researcher. All interviews were audio- and video-recorded and subsequently transcribed for analysis.

Students' cause-and-effect integrated essays were also evaluated by an EAP instructor from the same program using the same analytic rubric. The instructor had been teaching a different section of the same EAP course at the time of the study. In addition, she had contributed to the development and validation of the rubric. Thus, she was familiar with the test task instructions, prompts, source materials, and the rubric criteria. The scores from the instructor and the students were compared to identify similarities and differences in their perceptions.

Analysis

Following a case study methodology, data sources were analyzed taking an inductive approach to discover themes and patterns from unique cases (Creswell, 2013; Patton, 2002). An inductive approach emphasizes an initial examination of individual cases before combining those cases, and thus, the primary research focus was placed on the analysis of introspective interviews, and self-assessment from each student. For this purpose, self-efficacy questionnaire items were scored from 1 to 5 and then summed for each student following Abdel-Latif (2015). The interview data were analyzed qualitatively using open and axial coding (Corbin & Strauss, 2015). This process involved reading the transcripts recursively and making notes to identify initial codes for establishing tentative and provisional categories. Using the initial codes as a guide, axial coding was carried out to review, refine, and group the initial codes into more meaningful categories. Finally, the key themes extracted from individual cases were combined and included in the final report. All analyses were conducted by the researcher and then verified by a trained research assistant to ensure the veracity of findings.

Findings

Individual profiles below provide details about students' background, self-efficacy ratings, and general perspectives about integrated writing tasks. The first set of findings discussed after the student profiles pertains to students' conceptualization of classroom-based integrated writing tasks, which were discussed under three themes: a) task requirements, b) test conditions, and c) feedback from the instructor. These themes are individually described below and discussed in relation to students' test preparation, test taking strategies and the use of available support systems, such as attending office hours. The second set of findings concerns students' self-assessment of their own writing and engagement with the rubric categories. These findings were examined in relation to results from the textual analysis and instructor evaluation of the essays.

Student profiles

Liu – with Low Self-efficacy

Before immigrating to Canada 13 years ago, Liu had completed a master's degree in business administration. She was in middle school when she first started learning English in China. At the time of the study, it was her second semester as an undergraduate student in accountancy. Based on her performance on the placement test, she was required to take both writing courses (Course 1, and 2) in the EAP program. When asked to compare Course 1 and Course 2, Liu said: "This semester I feel a little bit more difficult compared to last semester. I am very nervous before the tests this year because I really have no idea how I could prepare". Her responses to the open-ended questions before the interview suggested she felt more confident in her receptive skills (reading and listening) compared with writing and speaking abilities. On the other hand, she remarked that her reading comprehension skills do not help her attain excellence in writing: "In my study, I am reading and reading and reading but it is not enough to write the main idea and construct essay [sic]". During our one-to-one meeting, Liu discussed her experience with integrated writing assessment focusing on her physical and cognitive reactions (e.g., feelings of panic, test anxiety) to assessment tasks. Her writing self-efficacy score was 56, which was the lowest among the three students in this study.

Qian – Questioning Test Purposes

Holding a bachelor's degree in business administration in China, Qian worked as a professional for 10 years and actively used English for work purposes in his home country. A year after moving to

Canada, he was admitted to the university for a bachelor's degree in accountancy. During his first semester as an undergraduate student, he was required to register in the EAP writing course (Course 2) along with two other courses in his degree program. While describing his experience with the coursework, Qian said: "So far, to tell you the truth, I did not think I gained much from the course". His writing self-efficacy score was 71, which is much higher compared to Liu, although he identified writing in English as one of his "weaknesses". The crux of our discussion focused on his challenges with incorporating information from source materials along with personal opinions. Qian also questioned test purposes in terms of assessing background knowledge, which seemed to distract him from engaging with course content and attaining higher scores in exams.

Moza – a Motivated Learner

Right after finishing high school in Lebanon, Moza moved to Canada to study Human Resource Management as an undergraduate student. In her first semester, apart from the EAP writing course (Course 2), she was registered in prerequisite courses, macroeconomics and math, offered by the business department. She reported speaking French as an L2, adding that she has focused on communicating in English since she moved to Canada, which helped her with the EAP coursework. Regarding her background in academic writing, she remarked that despite taking an academic writing course in Grade 12, she did not have to cite information from the sources. Even so, she acknowledged that the EAP writing course is not very challenging for her, which reflected on her self-efficacy score (82). Overall, Moza's experience with the assessment tasks in the EAP course seemed to be rather positive. She considered integrated writing as an important skill for university courses as she said: "I definitely think I need to do well in the exams to be successful in my other courses".

Keeping in mind the students' writing background, self-efficacy, and general perceptions of their writing course, the following section expands on their perceptions of integrated writing assessment by exploring their thoughts on the task requirements, and how the test conditions and instructor feedback impacts their task performance.

Student perceptions of integrated writing assessment

Task Requirements

An important concern related to task requirements raised by these students was moving from a summary task, elicited in the midterm exam, to writing an integrated cause-and-effect essay. Because summary writing does not require incorporating personal opinions, as different from integrated writing, the students felt disoriented since they were unsure how much of the information in their cause-and-effect essays should come from the source materials and whether they were allowed to discuss their personal opinions. For example, Liu reported preparing for the integrated writing test by reading the pre-assigned source materials carefully and then summarizing the main ideas in her own words so that she could cite them in the exam. She described her experience with the classroom-based cause-and-effect integrated writing test as follows:

So, I wrote about population in the first paragraph and then summary from the readings. Then, in the second paragraph, I also gave another summary but according to the teacher's requirement, maybe I should give my opinion in every paragraph. I did not know that.

Qian, on the other hand, was aware that he needed to develop personal arguments, but he was not

certain if his opinions needed to be validated by source-text information. His comment below illustrates his confusion in terms of establishing a balance between personal ideas and information from sources:

In the exam, I was not sure when I needed to cite. When I write my own opinion, it seems that it is not that supportive. If I know the point of view is from a source, I cite it. But sometimes I just express the same or similar opinions without reading the articles. I just expressed my own opinions, with no citation and maybe it was wrong. I do not know if you meet this kind of situation. So, what should we do? Am I wrong?

The written task instructions for the cause-and-effect essay included that students must use “a minimum of three sources to support their ideas” and they should write “a clear thesis statement that connects logically with the three supporting paragraphs.” After the midterm examination, which targeted summary writing, the students completed a pedagogical task, which required them to write a cause-and-effect essay using information from sources, and the course instructor provided them feedback on their responses using the current assessment criteria from the EAP program. They had also taken two ungraded quizzes, which targeted linguistic revision of source-text language, and accurate representation of source content. On the other hand, it appears from these comments that their particular challenges were associated with retrieving relevant information from the task environment, rather than integrating source information (Plakans, 2009). This perhaps encouraged their efforts to address the prompt within their own capacity.

In case of Moza, despite reporting that she had tried to “extract numbers, examples and important stuff” from the source materials while preparing for the test, she acknowledged struggling with organizing the discourse, as shown in this comment:

It is kind of hard to be inspired by the texts, but at the same time have your own idea. Like it is hard to separate them. So, you have just a bit here and here, and then you mix them in your own writing. That was my problem while writing this essay with time pressure.

Developmental issues, such as Moza's, which concern the application of knowledge from summary writing to integrated writing tasks could be assisted by extended instruction, and targeted practice. On the other hand, Liu and Qian's misconceptions point towards a problem in respect to task representation – a process that entails understanding the task instructions, establishing goals, and devising strategies to achieve these goals (Cheng, 2009; Plakans & Liao, 2018). Previously, researchers have shown that task type can impact L2 students' task representation, and textual borrowing strategies, such as the number of words borrowed from the source materials (Spivey, 1991; Shi, 2004). In addition, understanding and following task instructions have been found to predict students' integrated writing test performance (Plakans & Gebril, 2013). Therefore, students' insights into task instructions are useful to address the factors affecting test validity.

Test Conditions

Test conditions were discussed in relation to background knowledge and psychological factors by these students. First, they described their concerns about the high-stakes nature of classroom-based integrated writing tests. Timing, more specifically, contributed to a perception of increased anxiety and fatigue within testing conditions. For example, Liu, who commented: “We have 3 hours and 1 shot! The time is clicking, stressful and I feel I cannot organize my paragraph”. Similar comments were made by Moza, who believed that she could not perform well enough on the test due to time

pressure and feelings of exhaustion:

Honestly, I do not think I did any paraphrasing to the last example from the reading because after three hours, I was so tired, and I did not put all my focus on it. If I can write a second draft, definitely I would do better on this essay.

Qian's challenge, on the other hand, was tethered to not having an opportunity to use vocabulary and grammar creatively while focusing on other aspects of integrated writing, which made him question the test purposes:

I think grammar is my strong point, but I do not have a very large academic vocabulary. When writing this essay, I tried to use citations with good grammar and vocabulary, but I needed time for fixing other things like my topic sentence, so I cannot support my essay with enough ideas from readings. I do not know what exactly teacher expects from us.

Moza and Qian also noted that having no background information about the prompt increased their stress and yielded to lower performance on the test. Moza remarked that if she had prior knowledge about the topic, she could have done a better job with paraphrasing:

In the exam, we had to write about the environmental impact of bottled water on the environment, and I had no knowledge about this subject, so it was hard for me to understand the readings and use my own words when paraphrasing. In the first weeks, the topics were much better for me like nutrition and happiness. I was able to write better as I am familiar with these subjects.

In case of Qian, he was challenged by the lack of choice in topics as he expected the prompts to be more directly relevant to the pre-assigned readings:

What we need to do is to the first is to read the materials from the course-pack before the test. But the two topics teacher gave us to choose was different from what I read in the materials. So, that is another reason I did not do well enough in this test. Is it for testing my knowledge or ideas from other people's passages?

These perceived challenges suggest that timed exam conditions can be cognitively demanding for the students and impact their task performance. Integrated writing tasks have been shown to elicit more sophisticated linguistic and organizational features compared to independent writing (Cumming, 2013; Plakans & Gebriel, 2017). Additionally, these tasks require drawing on source materials for validating content in students' texts, which indicates that reading ability is an important underlying construct in integrated writing assessment (Leki & Carson, 1997; Watanabe, 2001). Because these students needed to divide their limited attentional resources to multiple subskills (e.g., generating personal ideas, organizing essay structure, linguistic revision of source ideas), exam conditions possibly limited their ability to demonstrate their learning from classroom activities. This was particularly evident in Liu's comments about the time pressure and Qian's questions about test purposes. Similar findings have been reported in DeLuca and Cheng's (2011) research, in which students perceived psychological factors such as anxiety and fatigue as detrimental to their test performance. Because assessment conditions tend to play a role in students' test appeal, and task performance (Davies et al., 1999), it is important for classroom instructors and test developers to minimize the impact of construct-irrelevant variance by obtaining evidence from test-takers.

Feedback from the Instructor

The students discussed the effect of feedback on their task performance in terms of understanding the course instructor's expectations and the usefulness of scoring criteria. In general, the students seemed to engage with teacher feedback; however, they held both positive and negative views of its usefulness. For example, Liu, among the three students had the most favorable opinion about teacher feedback:

I think to improvement for me is when I get the feedback from the professor on my writing. So, I know where I need to more to focus on all and where is my weakness. This is a great help to improve my writing.

When asked specifically about what types of feedback she had received from the instructor on the pedagogical tasks, Liu said:

The professor tells me to use the words and construct sentences in proper way. I think that she knows what I mean, what I want to express, but she wanted me to correct my language for better express my ideas and it takes time. Also, the professor wanted me talk about only one idea in a paragraph. I know I need more ideas.

As different from Liu, both Qian and Moza regarded feedback as discouraging and unclear, at times. For example, Qian commented:

I did not think I did so bad until I got it the feedback on the classroom cause-and-effect essay. But the result proved different from what I thought. The teacher said I cannot always grasp the key point. What I think is important in the reading passage, the teacher do not think so. I am not sure how I change my understanding of what is important and what is not.

Moza's experience, on the other hand was pertinent to not being able to address teacher feedback without making major changes in her essay. She expressed her concern saying: "When the teacher points out the mistakes, then I feel that I should change like my whole text, not just the mistakes." Elaborating on the kinds of feedback that she had received from the instructor, Moza said:

Like every time I get feedback from the teacher, she said "wording" both in summary and other writing assignments. I do not know what exactly "wording" is. I never asked her. I asked her so many things, but I never thought to ask what "wording" is. Maybe I tend to sophisticate my language and I add ideas and ideas, but at the end of the day, no, that is not good. Maybe I should ask her before the next test.

These varying perceptions about feedback suggest that it was not entirely clear to the students how they could address specific comments effectively from the instructor. Although the instructor held weekly office hours and encouraged students to attend, all three students remarked that they preferred asking their questions during regular class meetings. They also reported consulting with their classmates about the feedback and looking up sample essays online to improve their writing performance. Regarding the feedback they received on formative tasks (i.e., ungraded quizzes targeting source use features), the students agreed that it was useful to practice writing short texts and receiving feedback in a timely manner, which supports the assumption that they attended to ongoing feedback as part of learning and test preparation in this particular context.

Importantly, except the formative assessment, the feedback from the instructor generally focused on the generic aspects of writing in a second language rather than targeting students' source integration. Researchers have shown that L2 writers' challenges with integrated writing tasks mostly relate to understanding source materials, performing linguistic revision of source ideas and representation source content accurately (e.g., Keck, 2014; Plakans & Gebril, 2012, 2013; Wette, 2010). In the current study, despite her high self-efficacy, Moza identified paraphrasing and summarizing as the most challenging aspects of writing from source saying: "The hardest part is you have to rephrase it. You cannot just copy and paste it, and it's new to me". Similarly, Liu and Qian mentioned having no prior experience in source-based writing and experiencing difficulties in terms of selecting ideas from sources and linking them to their personal opinions. Although it is possible that the course instructor prioritized language use and idea generation over source use at this time of the semester, it is a key for source use to be emphasized as an integral part of corrective feedback in EAP settings to prepare students for writing in their disciplines.

Turning to the use of scoring criteria as part of teacher feedback, all three students agreed that it is beneficial to receive a score based on the rubric, which is used to evaluate their integrated writing performance in the exams. For example, Qian said: "When I saw my performance on the rubric, I kind of knew why I am not "above standard". Liu and Moza shared the same perception as Qian although none of the students had used the rubric for self-assessment or exam preparation. Also, when asked to comment on rubric categories and descriptors, the students could hardly recall those without actually seeing the rubric. Their perceptions were centered upon the performance levels in the current EAP program rubric (i.e., below standard, standard, above standard), rather than the constructs outlined in the criteria. Thus, the rubric appeared to function as a tool for students to reflect on their performance levels, rather than replacing or constituting corrective feedback. Because evaluation criteria promote autonomous learning and self-monitoring in L2 writing classrooms (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013), it is critical to make effective use of rubrics to facilitate students' understanding of teacher feedback.

Taken together, the students experience challenges with understanding and following test instructions, and they view test condition and feedback effects as a negative influence regarding validity of classroom-based integrated writing assessment. The following section will describe the students' use of an integrated writing rubric for self-assessment in connection with the results from textual analysis and instructor evaluation of the essays.

Self-assessment using an analytic rubric

During the one-on-one meetings, the students were introduced to the new assessment criteria, which evaluated source use as a separate category, and included more detailed and explicit descriptors for content and organization compared to the current EAP program rubric. They were encouraged to ask questions after the initial discussion of the rubric and while elaborating on their scoring decisions in reference to the specific descriptors. Even though they had not encountered the rubric prior to this study, the students did not seem to experience much difficulty understanding the descriptors. On the other hand, the way they treated the rubric during self-assessment showed variation. For example, Liu approached the criteria more holistically, compared to Qian and Moza and rather than highlighting specific descriptors across different score levels, she assigned a score of 2 (out of 4) for source use and language use, and 3 for content and organization. Her comments regarding her source use performance were as following:

Because source use, I think I am not perfect, not even 3. I know in this paragraph I needed another resource to support this idea. I try to write in my words but maybe not perfect. Also, my sentences and vocabulary should be better when summarizing.

As shown in Table 3, there was an overlap between Liu's self-assessment and the instructor evaluation of her essay using the same criteria although she received a slightly higher score for language use from the instructor.

Table 3 Comparison of self-assessment and the instructor evaluation

	Content		Organization		Source Use		Language Use	
	Self	Instructor	Self	Instructor	Self	Instructor	Self	Instructor
Liu	3	2.5	3	3	2	2	2	2.5
Qian	3	2.5	3.5	3	2	2.5	2.5	3
Moza	3.5	3	4	3.5	3	3	3.5	3.5

Textual analysis of Liu's essay for linguistic revision of source ideas revealed instances of verbatim copying and close paraphrasing in her essay (see Table 4). When asked about these particular instances, she commented:

The professor knows this [unacknowledged source use] paraphrasing is from the reading. So maybe I still need the author and date, right? ... I do not remember but I think I changed the words in this example [verbatim copying].

Table 4 Textual analysis

Source use variables	Liu	Qian	Moza
Essay length (N of words)	501	540	590
Citation density (N of citations)	5	4	6
Instances of verbatim copying (4+ strings)	2	0	1
Direct quotations	0	3	0
Close paraphrasing	1	0	1
Unacknowledged source use	1	0	1

As for Qian, he used the criteria more analytically, and annotated his essay based on the individual descriptors corresponding to his task performance. For example, he highlighted overreliance on quotes under source use and limited range and variety of vocabulary under language use and provided specific examples from his essay to support his scoring decisions. Qian regarded his textual organization "almost perfect", except for the use of transition words between the paragraphs. He was mostly critical of his source use, which, he thought, had a detrimental effect on content development:

I know I have deficiency in quoting and citing some information. So, in terms of this, I gave myself score 2. Also, there were few ideas supported with information from resources. This effects my content, as well.

Although Qian's self-assessment did not align well with the instructor scores (Table 3), he was able to use the criteria insightfully. He considered self-assessment as an eye-opening practice as shown by his comment below:

This rubric is very very useful because it gave me some guidance to judge my writing from another point of view. I need to do this more often and criticize myself from another perspective. It helps me a lot.

Turning to Moza, as a reflection of her high self-efficacy, she focused on the higher scores levels (3 and 4) while applying the rubric on her cause-and-effect writing. She was the least confident in her source use and assigned herself a score of 3. While explaining her scoring decisions, she pointed to the unacknowledged source use in the first body paragraph and commented:

In this paragraph, 30% of the information is from the reading, but the rest is from my background knowledge. Now looking at the criteria, I feel that I should have mentioned the source.

Moza acknowledged the instance of close paraphrasing in her essay, and she identified fatigue as the leading factor. Regarding the source verbatim copying in the introduction paragraph, she said: "I thought that since it is the hook and it is a general idea, it is not that important to cite the source." Overall, the instructor's evaluation of Moza's essay (Table 3) was consistent with her self-assessment. She attained the highest scores among the three students for all rubric categories. In terms of her self-assessment experience, Moza commented: "It is so good to grade yourself. I should do this more often."

The fact that students were able to use the evaluation criteria for self-assessment with ease suggests they could benefit from having access to a detailed and explicit integrated writing rubric in their EAP writing courses. All students regarded rubric use for self-evaluation as useful and their self-assessment showed similarity to the instructor evaluation. In particular, Qian and Moza were precise in their use of source use criteria as they were able to identify the potential problems related to linguistic revisions of source ideas. On the other hand, Liu and Moza shared the misconception with regards to the requirement for explicit citation of source ideas. This problem could be addressed by providing students explicit feedback on their source integration patterns. In addition, previous research which focused on L2 writers' integrated writing development have provided evidence of improvement in students' citation behavior after receiving targeted instruction (e.g., Hendricks & Quinn, 2000; Machili *et al.*, 2019). Thus, it is important to reconsider assessment and feedback as integral components of learning and instruction and incorporate students' voices in assessment decisions, including the development and refinement of evaluation criteria.

Discussion and Conclusions

Guided by L2 writers' experiences and challenges with classroom-based integrated writing tasks, the purpose of this study was to shed light on students' conceptualizations of integrated writing assessment in an EAP context and their use of an analytic rubric for self-assessment. Drawing on multiple data sources and adopting a case study approach, the study centralized a focus on students' contextualized challenges with integrated writing assessment. Three major themes emerged from the findings, which pertain to the students' conceptualizations of classroom-based assessment: task requirements, test conditions, and feedback from the instructors. First, although the students had practiced writing an integrated essay as part of pedagogical practices, they needed additional time and further instructions after the midterm examination to conceive of integrated writing task requirements. Their particular concerns were associated with the use of personal opinions along with the information from sources, and citation of the source materials. These concerns might stem from different individual characteristics, such as cultural and rhetorical background (Shi, 2004), or

contextual factors, including the types of pedagogical tasks students were exposed to prior to the test, as well as the nature of task instructions by the instructor (Greene, 1995; Plakans, 2009, 2010). The fact that these students were able to articulate their experiences with classroom-based assessment provides evidence that students' perspectives can enhance test processes and procedures.

Second, time pressure and lack of background knowledge about the topics prompted negative feelings, such as test anxiety, and fatigue, and affected the students' performance. As discussed by Cheng and DeLuca (2011), a decrease of test anxiety can motivate students to perform better on tests. In addition, allocating additional time for cognitive processes (e.g., planning, review) during the assessment helps students develop metacognitive strategies, which would guide them to monitor the writing process (Gebril & Plakans, 2009). Therefore, student perceptions of testing conditions are needed to ensure the validity of outcomes from test scores.

A final contextual assessment challenge that these participants encountered entailed the type of feedback that they received on their integrated essays. The students valued the feedback on the pedagogical tasks and ungraded quizzes; however, they required additional training to apply it to the integrated writing task. In some instances, they regarded the instructors' comments on their writing negatively, perhaps because of a lack of information on how to interpret the individualized feedback on their content and organization. This concern was accompanied by their approach to the current integrated writing rubric, which did not seem to facilitate their self-monitoring. These findings suggest construct-oriented rubrics need to be tailored for student use to facilitate their interpretation of test results (Alderson & Hamp-Lyons, 1996) and promote transparency in assessment (Crusan, 2015; Hudson, 2005).

Regarding the students' use of an analytic rubric for self-assessment, they were able to understand the constructs and identify the potential issues in their writing in light of the rubric descriptors. Integrated writing tasks in this study context requires students to connect source ideas with their personal opinions about the topic. The rubric students used for self-assessment targets source use as a separate category, and thus, the students had the opportunity to reflect on their source integration. Although they were mixed in their approach to the criteria, they deemed self-assessment as useful, and they were encouraged for purposing it as part of test preparation. This finding aligns with previous studies which suggested a link between the use of dialogic assessment methods and students' development of long-term motivation (Cheng & Fox, 2017; McMillan & Hearn, 2008). In addition, it contributes further evidence for the role of evaluation criteria in promoting autonomous learning and self-monitoring (Benson, 2007; Jonsson, 2014).

Implications and limitations

Overall, findings from this study suggest that student perceptions of fundamental aspects of language assessment can contribute to best practices in EAP contexts and offer practical considerations for stakeholders, including course instructors and test developers. First of all, because transitioning from summary writing to integrated writing tasks was challenging for these students, EAP instructors and curriculum developers might consider introducing students to response writing and providing them examples of contextualized source use to help bridge the gap between summary writing and integrated writing tasks. In addition, to facilitate students' task representation, instructors may need to flesh out the verbal and written task instructions. The use of practice tests supplemented with sample responses might familiarize students with the test format and sample prompts might also help students internalize the test requirements. Given that learning to write from sources is a gradual process of development, classroom-based assessment tasks might be designed taking a processed based approach. Completing a draft before writing the final version in a timed exam condition might

help with test-anxiety and fatigue associated with exam conditions. Furthermore, to develop students' autonomy, enhance motivation, and encourage their engagement with teacher feedback, it is essential to introduce them to revision strategies and to the use of evaluation criteria for self-assessment. Most importantly, integrated writing rubrics need to be localized through taking an evidence-based approach to rubric design. Evaluation criteria which represent the curricular objectives will address students' concerns regarding the interpretation of teacher feedback and centralize a focus on individual aspects of source use.

The present findings offer evidence of EAP students' challenges with understanding integrated writing task instructions and developing metacognitive strategies. These findings should be interpreted with caution because of some inherent limitations in the study design. First, students' perceptions were inevitably influenced by the local teaching and assessment practices, such as course delivery format (remote), the exam conditions, and the evaluation criteria. It is unclear whether the outcomes would be similar in different local assessment contexts. In addition, due to global pandemic, only three students from the EAP writing course volunteered to participate in this study. An important agenda for future research is to adopt a mixed-methods design and with more diverse student populations to account for the mediating factors within an integrated writing task environment. Finally, this study has not tapped into students' L1 literacy skills, which could impact their perceptions of classroom-based integrated writing assessment. To provide a better understanding of individual writers' challenges with integrated writing tasks, additional research that focuses on the relationship between students' L1 backgrounds and their test-taking strategies is needed.

References

- Abdel-Latif, M. (2015). Sources of L2 writing apprehension: A study of Egyptian university students. *Journal of Research in Reading*, 38, 194–212. <https://doi.org/10.1111/j.14679817.2012.01549.x>
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language testing*, 13(3), 280–297. <https://doi.org/10.1177/026553229601300304>
- American Psychological Association (2020). *Publication Manual of the American Psychological Association* (7th ed.). <https://www.apa.org/pubs/books/publication-manual-7th-edition>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Benson, P. (2007). Autonomy in language teaching and learning. *Language teaching*, 40(1), 21. <https://doi.org/10.1017/s0261444806003958>
- Butler, Y. G. (2016). Self-assessment of and for young learners' foreign language learning. In Nikolov, M. (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 291–315). Springer.
- Butler, Y. G. (2019). Assessment of young English learners in instructional settings. In Gao, A. (Ed.), *Second handbook of English language teaching* (pp. 477–495). Springer.
- Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing*. Advanced online publication. <https://doi.org/10.1177/0265532221992274>
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study* (Vol. 21). Cambridge University Press.
- Cheng, F. W. (2009). Task Representation and Text Construction in Reading-to-Write. *Journal of Pan-Pacific Association of Applied Linguistics*, 13(2), 1–21.

- Cheng, L., DeLuca, C. (2011). Voices from test-takers: Further evidence for language assessment validation and use. *Educational Assessment*, 16(2), 104–122. <https://doi.org/10.1080/10627197.2011.584042>
- Corbin, J., & Strauss, A. (2015). *Basics of qualitative research*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Crusan, D. (2010). *Assessment in the second language writing classroom*. University of Michigan Press.
- Crusan, D. J. (2015). Dance, ten; looks, three: Why rubrics matter. *Assessing Writing*, 26, 1. <https://doi.org/10.1016/j.asw.2015.08.002>
- Cumming, A., Kantor, R., Baba, K., Erdoosy, U., Eouanzoui, K., & James, M. (2006). Analysis of discourse features and verification of scoring levels for independent and integrated tasks for the new TOEFL. *TOEFL Monograph No. MS-30*. Princeton, NJ: Educational Testing Service
- Cumming, A., Kantor, R. and Powers, D.E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal*, 86, 67–96. <https://doi.org/10.1111/1540-4781.00137>
- Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, 10(1), 1–8. <https://doi.org/10.1080/15434303.2011.622016>
- Davies, A, Brown, A, Elder, C, Hill, K, Lumley, T, & McNamara, T. (1999). *Dictionary of language testing*. Cambridge: Cambridge University Press.
- Erickson, G., & Gustafsson, J. (2005). Some European students' and teachers' views on language testing and assessment. A report on a questionnaire survey. *European Association of Language Testing and Assessment*. Retrieved from <http://www.ealta.eu.org/documents/resources/enlta%20activity%204%20report.pdf>
- Ferris, D. R., & Hedgcock, J. S. (2014). *Teaching L2 composition: Purpose, process, and practice* (3rd ed.). New York: Routledge.
- Gebriel, 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(1), 47–84.
- Greene, S. (1995). Making sense of my own ideas: The problems of authorship in a beginning writing classroom. *Written Communication*, 12(2), 186–218. <https://doi.org/10.1177/0741088395012002002>
- Hamp-Lyons, L. (2014). Writing assessment in global context. *Research in the Teaching of English*, 48(3), 353–362.
- Harwood, N., & Petri, B. (2012). Performance in the citing behavior of two student writers. *Written Communication*, 29, 55–103. <https://doi.org/10.1177/0741088311424133>
- Hendricks, M., & Quinn, L. (2000). Teaching referencing as an introduction to epistemological empowerment. *Teaching in Higher Education*, 5(4), 447–457. <https://doi.org/10.1080/713699175>
- Hirvela, A., & Du, Q. (2013). “Why am I paraphrasing?”: Undergraduate ESL writers' engagement with source-based academic writing and reading. *Journal of English for Academic Purposes*, 12(2), 87–98. <https://doi.org/10.1016/j.jeap.2012.11.005>
- Hudson, T. (2005). Trends in assessment scales and criterion-referenced language assessment. *Annual Review of Applied Linguistics*, 25, 205–227. <https://doi.org/10.1017/S0267190505000115>
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840–852. <https://doi.org/10.1080/02602938.2013.875117>

- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- 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>
- Knoch, U., Read, J., & von Randow, J. (2007). Re-training writing raters online: How does it compare with face-to-face training?. *Assessing writing*, 12(1), 26–43. <https://doi.org/10.1016/j.asw.2007.04.001>
- Leki, I., & Carson, J. (1997). “Completely different worlds”: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), 39–69. <https://doi.org/10.2307/3587974>
- Li, Y., & Casanave, C. P. (2012). Two first-year students’ strategies for writing from sources: Patchwriting or plagiarism? *Journal of Second Language Writing*, 21(2), 165–180. <https://doi.org/10.1016/j.jslw.2012.03.002>
- Llosa, L., & Malone, M. E. (2017). Student and instructor perceptions of writing tasks and performance on TOEFL iBT versus university writing courses. *Assessing Writing*, 34, 88–99. <https://doi.org/10.1016/j.asw.2017.09.004>
- Machili, I., Papadopoulou, I., & Kantaridou, Z. (2019). Effect of strategy instruction on EFL students' video-mediated integrated writing performance. *Journal of Second Language Writing*, 48, 100708. <https://doi.org/10.1016/j.jslw.2019.100708>
- Malone, M. E., & Montee, M. (2014). Stakeholders’ beliefs about the TOEFL iBT® test as a measure of academic language ability. *ETS research report No. RR-4-42*. Princeton, NJ: Educational Testing Service.
- McDonough, K., Neumann, H., & Leu, S. (2018). *Concordia Written English Academic Texts (CWEAT) corpus*. Concordia University: Montreal, QC.
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational research review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Patton, Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Plakans, L., & Liao, J. T. (2018). Task representation. In *The TESOL Encyclopedia of English Language Teaching*. <https://doi.org/10.1002/9781118784235.eelt0532>
- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, 17(1), 18–34. <https://doi.org/10.1016/j.asw.2011.09.002>
- Plakans, L., & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, 22(3), 217–230. <https://doi.org/10.1016/j.jslw.2013.02.003>
- Plakans, L. (2009). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26(4), 561–587. <https://doi.org/10.1177/0265532209340192>
- Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, 44(1), 185–194. <https://doi.org/10.5054/tq.2010.215251>
- Plakans, L., & Gebril, A. (2017). Exploring the relationship of organization and connection with scores in integrated writing assessment. *Assessing Writing*, 31, 98–112. <https://doi.org/10.1016/j.asw.2016.08.005>
- Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: a learning-focused interpretation 1. *International Journal of Applied Linguistics*, 16(2), 163–188. <https://doi.org/10.1111/j.1473-4192.2006.00112.x>
- Rosenfeld, M., Leung, S., & Oltman, P. (2001). The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels. *TOEFL monograph series 21*. Princeton, NJ: Educational Testing Service.

- Sato, T., Ikeda, N. (2015). Test-taker perception of what test items measure: A potential impact of face validity on student learning. *Language Testing in Asia*, 5(1), 1–15. <https://doi.org/10.1186/s40468-015-0019-z>
- Shi, L. (2004). Textual borrowing in second-language writing. *Written communication*, 21(2), 171–200. <https://doi.org/10.1177/0741088303262846>
- Simons, H. (2009). *Case study research in practice*. Thousand Oaks, CA: Sage.
- Spivey, N. N. (1991). The shaping of meaning: Options in writing the comparison. *Research in the Teaching of English*, 25(4), 390–418.
- Uludag, P., & McDonough, K. (2022). Validating a rubric for assessing integrated writing in an EAP context. *Assessing Writing*, 52, 100609. <https://doi.org/10.1016/j.asw.2022.100609>
- Vlanti, S. (2012). Assessment practices in the English language classroom of Greek junior high school. *Research Papers in Language Teaching and Learning*, 3(1), 92–122. Retrieved from <https://rpltl.eap.gr/images/2012/RPLTL-03-01-fulltext.pdf>
- Watanabe, Y. (2001). *Read-to-write tasks for the assessment of second language academic writing skills: Investigating text features and rater reactions*. Unpublished doctoral dissertation. University of Hawaii at Manoa.
- Winke, P., Lee, S., Yoon, H.-J., Ahn, J. I., Choi, I., Cui, Y. (2018). The cognitive validity of child English-language tests: What young language learners and their native-speaking peers can reveal. *TESOL Quarterly*, 52(2), 274–303. <https://doi.org/10.1002/tesq.396>
- Robert K. Yin. (2014). *Case study research design and methods* (5th ed.). Thousand Oaks, CA: Sage
- Wette, R. (2010). Evaluating student learning in a university-level EAP unit on writing using sources. *Journal of Second Language Writing*, 19(3), 158–177. <https://doi.org/10.1016/j.jslw.2010.06.002>
- Yin, R. (2014). *Case study research design and methods* (5th ed.). Thousand Oaks, CA: Sage.
- Zhu, W. (2005). Source articles as scaffolds in reading to write: The case of a Chinese student writing in English. *Journal of Asian Pacific Communication*, 15(1), 129–152. <https://doi.org/10.1075/japc.15.1.09z>

Author biodata

Pakize Uludag is an Assistant Professor in technical communication in the Centre for Engineering in Society at Concordia University. Her interests include academic writing, language assessment and corpus linguistics.

Appendix A: Questionnaire Items

PART 1: English writing self-efficacy scale

Directions: Below are some statements about your English writing. There are no right or wrong answers to these statements. Please indicate the degree to which each statement applies to you by ticking to what extent you strongly agree (5) or strongly disagree (1).

	Strongly agree Strongly disagree				
	5	4	3	2	1
I am not good at writing in English.	5	4	3	2	1
It is easy for me to write good essays in English.	5	4	3	2	1
When I hand in an English essay I know I am going to do poorly.	5	4	3	2	1
I expect to do poorly in English writing classes even before I enter them.	5	4	3	2	1
I feel confident in my ability to clearly express my ideas when writing in English.	5	4	3	2	1
People seem to like what I write in English.	5	4	3	2	1
I do not think I write in English as well as my classmates.	5	4	3	2	1
When my class is asked to write an essay, mine is one of the best.	5	4	3	2	1

On a scale from 5 (very confident) to 1 (very unconfident), how confident are you that you can perform each of the following English writing skills?

	Very confident Very unconfident				
	5	4	3	2	1
Correctly spell all the words in a one-page essay.	5	4	3	2	1
Correctly punctuate a one-page essay.	5	4	3	2	1
Correctly use parts of speech (i.e., nouns, verbs, adjectives, etc.).	5	4	3	2	1
Write simple sentences with proper punctuation and grammatical structure.	5	4	3	2	1
Correctly use plurals, verb tenses, prefixes and suffixes.	5	4	3	2	1

Write an essay with appropriate vocabulary.	5	4	3	2	1
Write compound and complex sentences with proper punctuation and grammatical structure.	5	4	3	2	1
Write a strong paragraph that has a good topic sentence or main idea.	5	4	3	2	1
Organize sentences into a paragraph so as to clearly express a theme.	5	4	3	2	1
Write an essay with a good overall organization (i.e., ideas in order, effective transition, etc.).	5	4	3	2	1

(Adopted from Abdel Latif, 2015)

PART 2: Open-ended Questions

Please answer the following open-ended questions providing details about experience in your English writing course this semester:

1. What types of writing have you done/ are you currently doing in your English writing course where you used information from source materials (readings/articles)?
 2. What are the things you find difficult in your English writing course this semester?
 3. Which specific activities have you found useful to develop your writing skills in your English writing course?
 4. What do you think you do well when writing an essay in your English writing course?
 5. What do you struggle with when writing an essay in your English writing course?
 6. Please discuss how difficult/easy it is for you to read and understand the ideas in English academic reading texts (for example, the articles/readings in the course pack).
 7. Please discuss how difficult/easy it is for you to find good examples from the readings/articles in the writing course pack to support an idea you have in your text.
 8. Please discuss how difficult/easy it is for you to select information from the readings that you have read to include in your own text.
 9. Please discuss how difficult/easy it is for you to use information from the readings when you are writing an essay in your writing course.
- Please discuss how difficult/easy it is for you to use correct punctuation and formatting (APA Style) to show where the information came from.

Appendix B: Integrated Writing Rubric

	Content	Organization	Source Use	Language and Mechanics
4	<ul style="list-style-type: none"> - The prompt is fully addressed; the ideas are explored in-depth with supporting details. - The student's position is clear in the essay, and personal claims are always supported with evidence from sources. - Acknowledgement of and response to opposing view(s) is thorough and effective. 	<ul style="list-style-type: none"> - The essay has a clear and focussed thesis; introduction and conclusion are complete and effective. - Topic sentences clearly stem from the thesis and body paragraphs are cohesive. - Ideas between and within paragraphs are linked with smooth and effective transitions. 	<ul style="list-style-type: none"> - Information from sources is always relevant and accurate. - Source information is paraphrased/summarized in the student's words with structural and lexical changes. - Source information is cited properly. 	<ul style="list-style-type: none"> - The essay includes a variety of sentence styles and length with no major structural and language problems. - An extensive range and variety of vocabulary is used accurately. - The essay is virtually free of punctuation, capitalization, and spelling errors.
3	<ul style="list-style-type: none"> - The prompt is addressed adequately; the ideas are explored with supporting details. - The student's position is generally evident in the essay, and personal claims are mostly supported with evidence from sources. - Acknowledgment of and response to opposing view(s) is mostly effective. 	<ul style="list-style-type: none"> - The thesis is mostly clear; introduction and conclusion are mainly complete. - Topic sentences generally stem from the thesis and body paragraphs are mostly cohesive. - Most ideas between and within paragraphs are linked with transitions. 	<ul style="list-style-type: none"> - Information from sources is mostly relevant, and accurate but not quite enough/a bit too much. - Source information is mostly paraphrased/summarized in the student's words with adequate structural and lexical changes. - Source information is mostly cited properly. 	<ul style="list-style-type: none"> - The essay includes an appropriate range of sentence styles and length with occasional structural and/or language problems. - An adequate range and variety of vocabulary is used mostly accurately. - There are occasional spelling, capitalization, and punctuation errors, which do not interfere with meaning.

2	<ul style="list-style-type: none"> - There is some attempt to address the prompt; supporting details were used occasionally. - The student's position is somewhat evident in the essay; opinions are hard to distinguish from source information OR only few arguments are supported with evidence from sources. - Acknowledgment of opposing view(s) might be present but unclear/insufficient. 	<ul style="list-style-type: none"> - The thesis is present but unclear; introduction and conclusion are attempted but lack some required elements. - Inappropriate topic sentences may be present; body paragraphs show structure but lack cohesion. - Some ideas between and within paragraphs may be linked with transitions. 	<ul style="list-style-type: none"> - Information from sources may be irrelevant/inaccurate; there is some evidence but not enough. - Source information is somewhat paraphrased/summarized but too similar to original; there might be some unacknowledged information from sources. - There are a number of problems with citation of source information. 	<ul style="list-style-type: none"> - Sentences show structural and/or language errors and little variety. - Limited range and variety of vocabulary is used with some accuracy. - There are several errors of spelling, capitalization, and punctuation errors, which may interfere with meaning in places.
1	<ul style="list-style-type: none"> - The prompt is minimally addressed; the ideas are underdeveloped. - The student's position is rarely evident in the essay; there is no sense of individuality/no evidence is used to support arguments. - Opposing view(s) are not acknowledged. 	<ul style="list-style-type: none"> - There is no discernible thesis; introduction and conclusion are missing/ lacking required elements. - Body paragraphs lack structure and cohesion. - Transitions are used minimally. 	<ul style="list-style-type: none"> - Little or poor source evidence is given; the student relies mostly on quotes. - There are largely unchanged text chunks from sources/ much unacknowledged information suggesting plagiarism. - There are serious problems with citation of source information. 	<ul style="list-style-type: none"> - Sentences show multiple and serious structural and language errors and no variety. - Very narrow range of vocabulary is used with pervasive errors. - There are frequent errors of spelling, capitalization, and punctuation; Little, if any, proofreading is evident.
0	<ul style="list-style-type: none"> - The essay shows no engagement with the topic. 	<ul style="list-style-type: none"> - There is no indication of paragraphing. 	<ul style="list-style-type: none"> - No evidence is provided; no academic referencing/citation conventions are followed. 	<ul style="list-style-type: none"> - Intrusive and/or inaccurate language use which greatly impedes communication.