International Journal of Assessment Tools in Education
2026, Vol. 13, No. 1, 145–185
https://doi.org/10.21449/ijate.1641764
journal homepage: https://dergipark.org.tr/en/pub/ijate *Research Article*

# Training EFL writers in self-assessment of source use: Why it fell short

**Fang Wang**[1*]

[1]Arizona State University, Mary Lou Fulton Teachers College

**ARTICLE HISTORY**
*Received: Feb. 17, 2025*
*Accepted: Sep. 23, 2025*

**Keywords:**
Integrated skills assessment,
Source use,
Self-assessment,
EFL learners,
Secondary school.

**Abstract:** Self-assessment has attracted growing attention in language testing for its many benefits, such as fostering reflection, enhancing autonomy, and boosting confidence. This study investigated the effectiveness of training L2 writers to become proficient self-assessors of their source use ability. Ten secondary school EFL writers in China, ranging from beginner to high-intermediate levels, completed three integrated writing tasks, followed by self-assessment for each essay. Afterward, they received instructor assessment and feedback before submitting revisions for each task. A post-task interview was conducted to gather their perceptions of the training and self-assessment. Findings revealed that while some participants improved in source use, many struggled to benefit from self-assessment because they were confused about the evaluation criteria. For classroom practice, the study suggests that careful design of assessment tools, explicit modeling, scaffolding, and sustained teacher involvement are crucial for enabling students to engage meaningfully in self-assessment, thus promoting learning. For future research, this highlights the need to involve trained students, incorporate face-to-face interactions, use longitudinal designs, and explore how self-assessment can be adapted for different proficiency levels and how bilingual or first-language support may enhance its effectiveness.

## INTRODUCTION

In evaluating human cognitive abilities, particularly language learning, two perspectives are often distinguished: the internal (intrinsic) perspective and the external (extrinsic) perspective (Oscarson, 2013). The internal perspective pertains to learners' self-assessment, while the external perspective involves evaluation conducted by an external party. Self-assessment, widely recognized as an innovative educational practice, has been extensively utilized in language learning classrooms and rigorously studied in language testing research due to its strong correlation with effective teaching and learning outcomes.

The development of self-assessment has a long history. Documented practices appeared as early as the 1930s (Schuessler, 2010), with early research exploring the correlation between student self-ratings and teacher evaluations. Interest grew in the 1970s, partly due to efforts to redistribute power in education by engaging learners in assessment. By the late 20th century, self-assessment was also recognized in professional contexts as a valuable workplace skill. Today, it continues to attract attention in both academic and professional domains.

***Corresponding author at:** Fang WANG fwang117@asu.edu Arizona State University, Mary Lou Fulton Teachers College



**e-ISSN: 2148-7456**

Self-assessment benefits learners by clarifying learning objectives, encouraging engagement with criteria (Carless, 2007), and cultivating autonomy, confidence, and self-evaluation skills (Boud, 2000; Falchikov, 2006). For these reasons, it is now regarded as a cornerstone of learning-oriented assessment (LOA). Carless *et al*. (2006a, 2006b, 2015) introduced LOA to emphasize assessment processes that prioritize learning as well as measurement.

Research on self-assessment has generally focused on three key areas: reliability, validity, and practicality (Ross, 2006). Studies consistently show strong internal consistency across tasks, only partial alignment with teacher ratings, and positive impacts on learner achievement and motivation. Importantly, training learners to assess more effectively and involving teachers in the process enhances the benefits while reducing potential drawbacks.

Accuracy in self-assessment also depends on learner characteristics such as language proficiency, self-esteem, and motivation (AlFallay, 2004; Heilenman, 1990; Suzuki, 2015), as well as the design of assessment items. Items tied directly to task objectives or specific achievements tend to yield more accurate self-ratings (Butler & Lee, 2006; Ross, 1998). Factors such as item wording (Heilenman, 1990) and language of assessment (Oscarson, 1997) further influence results.

Although self-assessment and self-evaluation are often used interchangeably (Alverno College Faculty, 2000), they differ in scope: self-assessment typically refers to evaluating a single performance, while self-evaluation involves reviewing performance over time. Both can be integrated through portfolios, which provide longitudinal insights into learner development (Biggs, 1996; Hauser *et al*., 2005). Rolheiser (1996) highlighted how self-evaluation fosters a positive learning cycle of goal-setting, effort, achievement, judgement, and reflection (see Figure 1 below).

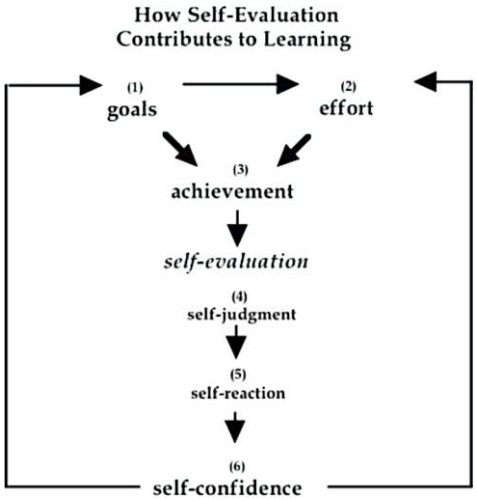



**Figure 1.** *Rolheiser's (1996) model of self-evaluation.*

Rolheiser and Ross (2001) later proposed a four-stage model for training learners to self-assess, emphasizing defining criteria, applying them with support, receiving feedback, and setting new goals (see Figure 2 below). These models underscore the importance of structured teacher guidance in cultivating effective self-assessment practices. Such structured approaches become particularly relevant in language testing, where integrated reading-to-write tasks have gained prominence for their authenticity in simulating academic writing (Plakans, 2008, 2009b). Because these tasks require learners to synthesize information across skills, they create rich opportunities for applying self-assessment cycles of goal-setting, evaluation, and reflection. Research has shown that integrated tasks reveal more about learners' strategic competence and source use than independent tasks (Cumming *et al*., 2016; Weigle, 2004). They also align with current test validity frameworks, which emphasize authenticity, cognitive processes, and learning potential (Purpura, 2016; Weir, 2005).

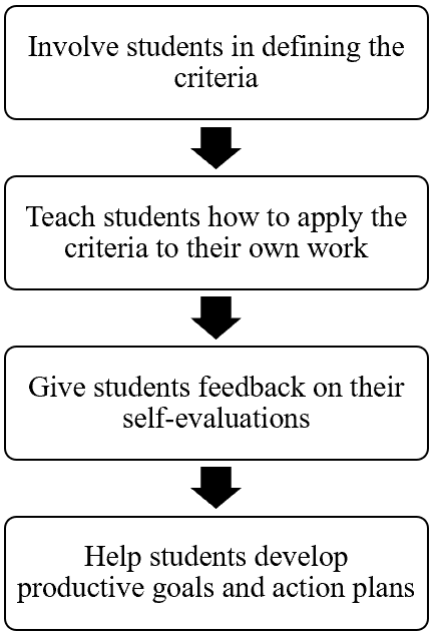



**Figure 2.** *Rolheiser and Ross's (2001) four-stage model for teaching self-assessment to learners.*

A critical challenge in integrated writing is the responsible use of source texts. Studies have identified common difficulties such as patchwriting, plagiarism, and challenges in paraphrasing (Keck, 2006; Plakans & Gebril, 2013; Shi, 2010). Self-assessment can help learners monitor their source use, reflect on appropriateness, and set goals for improvement. When combined with clear rubrics and teacher scaffolding, self-assessment not only enhances writing quality but also supports academic integrity (Gebril & Plakans, 2016). Supporting this perspective, Zhang *et al*. (2022) developed and validated a scale measuring L2 college learners' self-efficacy in reading-to-write tasks within an EFL context, identifying three dimensions: discourse synthesis self-efficacy, self-regulatory efficacy, and writing conventions self-efficacy. Their findings showed that self-regulatory efficacy and discourse synthesis self-efficacy significantly predicted performance on integrated writing tasks, highlighting the role of self-assessment in fostering learner confidence and effectiveness.

Despite these benefits, most research to date has focused on adult English language learners in English for Academic Purposes (EAP) contexts, leaving a gap in understanding how secondary-level EFL learners develop self-assessment practices for source use in integrated writing tasks. This study addresses that gap by examining the effectiveness of training secondary school EFL writers to self-assess their source use in integrated writing tasks within a workshop context. In doing so, it contributes to the literature on self-assessment and integrated writing while providing practical strategies to help young learners become more effective and autonomous evaluators of their own writing.

## METHOD

### Setting and Participants

This study took place at a bilingual school in Shanghai, China, where students were preparing to pursue undergraduate studies overseas. The school followed a U.S. high school curriculum that included academic subjects (e.g., English Literature, English Composition, Algebra, Chemistry, American / World History, World Geography, and Chinese Language Arts) alongside test preparation courses (TOEFL iBT, SAT, AP), as reported on the school's profile (https://www.jinpingguo.com/schoolProfile.aspx). Some students also attended cram schools outside regular hours to prepare for these exams. At this bilingual school, instruction was delivered in English by both native Chinese-speaking and native English-speaking teachers, all

of whom held at least a bachelor's degree in language or language education. Students who fulfilled the required coursework earned an American high school diploma.

After receiving IRB approval from my institution, I forwarded the recruitment flyer to the school principal, who helped distribute it to the Grade 11 and Grade 12 teachers, who then shared it with students and their legal guardians. These two grades were targeted because students at these levels are more likely to be familiar with integrated writing tasks, as they are preparing for English language proficiency exams. Ten Grade 11 and Grade 12 secondary EFL writers with intermediate (shown as "Independent user" in the table) or lower (shown as "Basic user" in the table) English language proficiency levels took part in the study (see Table 1 for their demographic information). It is important to note that participants' proficiency levels were interpreted using the Common European Framework of Reference for Languages to ensure consistent interpretation of scores.

**Table 1.** *Participant demographic information.*

| ID | Gender | Grade | TOEFL / IELTS score | Proficiency level | Part score (reading) | Part score (writing) |
|---|---|---|---|---|---|---|
| 1 | M | 11 | 3 | Basic user | 3 | 3 |
| 2 | M | 11 | 76 (or 6 in IELTS) | Independent user | 19 (or 6.5 in IELTS) | 18 (or 5.5 in IELTS) |
| 3 | F | 11 | 65 (or 6 in IELTS) | Independent user | 20 (or 6.5 in IELTS) | 16 (or 5 in IELTS) |
| 4 | F | 11 | 5 | Independent user | 5 | 5 |
| 5 | M | 11 | 4.5 | Independent user | 3.5 | 5.5 |
| 6 | F | 11 | 5 | Independent user | 5 | 5 |
| 7 | F | 11 | 50 (or 5.5 in IELTS) | Independent user | 16 (or 6 in IELTS) | 18 (or 5.5 in IELTS) |
| 8 | M | 11 | 4.5 | Independent user | 5 | 4 |
| 9 | F | 12 | 4.5 | Independent user | 5 | 4.5 |
| 10 | F | 12 | 4.5 | Independent user | 4 | 5 |

## Instrumentation

### *Integrated writing tasks*

Throughout the study, participants completed three integrated writing tasks. Each task required them to read two 200–300-word academic sources and then write a 250-word argumentative essay based on the information provided in the readings. Each task also included three additional components: task requirements and evaluation criteria, a writing prompt, and writing instructions. Together with the reading texts, these four sections comprised each writing task.

The task requirements specified that participants should read about two sides of an issue and develop a persuasive argument supporting their position. Argumentative tasks were selected because they typically require students to synthesize information and build arguments from materials provided, reflecting authentic academic contexts (Jamieson *et al.*, 2008). Additionally, participants were required to cite at least three points from the readings, and each essay should be approximately 250 words in length. The evaluation criteria focused on how effectively participants incorporated information from the readings and referenced the author(s) accurately. A standardized writing prompt followed the requirements and criteria: "Some people support [xxx], while others cast doubt on it. What is your opinion? Use reasons and/or examples to support your writing."

After the prompt, I provided several instructions to assist participants in their writing, adapted from the sample reading-writing task in Plakans and Gebril (2015). I also informed participants

that they could use any linguistic tools while writing, as long as they refrained from seeking external help or plagiarizing from other sources. Appendices A and B provide samples of the reading-to-write tasks developed for this study and the online essay paper on which participants wrote their essays, respectively.

### *Assessment checklist*

Wright (2019) suggests that a checklist can be an effective tool for self-assessment, especially during the editing stage of the writing process. In addition to its flexibility in being tailored to specific assessment purposes, a checklist is simpler than a rubric, which typically includes multiple score scales and performance descriptors for each level. As such, a checklist is more "learner-friendly" for L2 writers when conducting self-assessment and receiving feedback from instructors. Given these considerations, this study utilized a checklist focused solely on integration styles (see Appendix C for the assessment checklist).

To develop the checklist, I drew on Plakans and Gebril (2013) and Lowry's (2016) Choosing *& Using Sources: A Guide to Academic Research*. Plakans and Gebril categorized each T-unit in participants' essays that was associated with a source idea as either explicit (quoting or direct copying without quotation marks) or implicit (paraphrasing or summarizing) source use. In creating the checklist, I based the assessment criteria on three major categories of integration styles: direct quotation, paraphrasing, and summarizing. Numerous scholars have provided definitions for paraphrasing and summarizing. For instance, Hirvela and Du (2013) defined summarizing as condensing source information by retaining only the main idea, while paraphrasing involves maintaining a similar level of detail to the original text. These definitions clearly distinguish between the two types of source use by the level of specificity in incorporating source material. I made a slight adjustment to the definition of direct quotations, as participants in this study were provided with the authors' information in the source text. Table 2 lists the definitions.

**Table 2.** *Definitions of direct quotation, paraphrase, and summary.*

| | |
|---|---|
| Direct quotation | Verbatim language from source text set off by quotation marks, author's name, year, and page number. |
| Paraphrase | The writer paraphrases one idea from the source. |
| Summary | The writer summarizes or makes a generalization about two or more ideas from the source text. |

The next step was to clarify each criterion. I adapted the term "direct copying without quotation marks" from Plakans and Gebril (2013) and Cumming *et al*. (2005) to "properly referring to and citing the author(s)," as a descriptor under the direct quotation category. To further define paraphrasing and summarizing, I referred to Lowry (2016), which offers detailed guidelines on how to properly quote, paraphrase, and summarize sources.

In addition to the checklist, a comment section was included at the bottom of each criterion section. As the assessor of participants' essays, I provided feedback in this section to explain the rationale behind the performance evaluation (Andrade & Evans, 2012). Participants were also given the option to leave comments for themselves, though this was not mandatory, as providing constructive feedback requires extensive training (Min, 2005; Rahimi, 2013).

It is important to note that the assessment checklist was designed in English for two main reasons. First, I wrote the evaluation criteria in plain language to ensure clarity. Second, I explained these criteria in Chinese during the assessment training (discussed later) to ensure that participants fully understood how their essays would be evaluated.

### *Post-task interview*

The post-task interview was designed to explore participants' perceptions of the effectiveness of self-assessment in improving their subsequent writing. The interview consisted of two sections. The first section collected demographic information from participants, such as their grade level, first language, and most recent language test scores. The second section included questions focused on the research question, specifically aimed at understanding how participants perceived the self-assessment process (for interview questions, please see Appendix D). The results from the interview helped identify the key factors that influenced whether this novel approach was successful in promoting the learning of EFL writers.

## Data Collection

The study was conducted in 2021, during the COVID-19 pandemic. Prioritizing the health and safety of participants, and in accordance with the university's IRB regulatory and policy updates, face-to-face data collection was transitioned to online formats to minimize in-person gatherings (Office of the Vice President for Research, 2020). As a result, all data collection activities were carried out online. The following section will provide a detailed description of the main data collection venue, the writing workshop, which included an introductory meeting and process writing activities.

### *Writing workshop*

**The Introductory Meeting.** The workshop began with a one-hour introductory meeting with each participant. During this session, I introduced the study consent form, the integrated writing tasks, source use, the assessment checklist, and allowed participants to practice using the checklist to evaluate one essay. I also provided an overview of the three integration styles—direct quotes, paraphrasing, and summarizing—that participants would use in their writing.

Following part of Liu and Li's (2014) assessment training procedure, participants were introduced to the checklist that would be used to evaluate their work. I explained each section of the checklist to ensure that participants had the same understanding of the criteria. They were encouraged to ask questions and share any thoughts or concerns about the checklist.

Next, participants assessed two sample essays: an anchor essay (a sample essay response to the writing task) and another containing plagiarism and excessive quotations. Participants used the checklist to assess both essays, which helped them better understand the criteria listed in the assessment checklist. All explanations were provided in Chinese to ensure clear understanding during the meeting.

The introduction and assessment training materials were prepared in PowerPoint slides (see Appendix E). During the online meeting, I shared my screen with the participants. Although the training session was conducted online rather than in person, which is typically considered the conventional method for rating training, the delivery modality did not impact on the overall effectiveness of the training (Knoch *et al*., 2007). The PowerPoint slides were also provided to participants after the session for any additional review they might need.

**Process Writing.** After the meeting, the workshop lasted for about one month. Each week, participants read one assigned reading and wrote an argumentative essay. Then, they participated in self-assessment and sent me their results. Participants only received my feedback after they sent me their self-assessment results to avoid their judgement being affected by my evaluation. After collecting evaluations from participants themselves and me, participants were asked to revise their writing tasks and resubmit the updated essays. Participants were given about one week to complete each writing and self-assessment task to ensure the quality of the data. I assessed their essays using the same checklist and returned the revised essays to them.

As mentioned earlier, participants engaged in self-assessment and received instructor assessment and feedback each time they completed an integrated writing task. This combined use of instructor feedback and learner self-assessment is referred to as co-assessment by Dochy

*et al*. (1999). When reviewing participants' essays, I highlighted problematic language in bold and provided comments in angle brackets next to the relevant text. While using the checklist for evaluation, I ticked the checkbox to see if the essay met a specific criterion and crossed it out if it did not. To guide my feedback, I followed the feedback theories and sample feedback provided by Hattie and Timperley (2007). Additionally, I adhered to Lee's (2007) guidelines to ensure that my comments were consistent and concise, aiming to enhance learner understanding. All feedback provided in both the checklists and participants' essays was written in English, and no scores or grades were included as part of my assessment.

To ensure the reliability of my assessments and feedback, I collaborated with a writing tutor who had extensive experience working with L2 writers on academic writing. The tutor and I independently evaluated one participant's essay and then compared our assessments and feedback. After reaching agreement, I proceeded to assess the remaining essays and provide feedback on my own.

**Post-task Interview.** At the end of the study period, participants were invited to a post-task interview. The interview was conducted in Chinese based on participants' preference. All interviews were conducted online via Zoom. I used the Zoom recording function to record only the voice data. Figure 3 illustrates an overall schedule of the entire study. To thank participants for taking part in the study, everyone received $150 compensation.

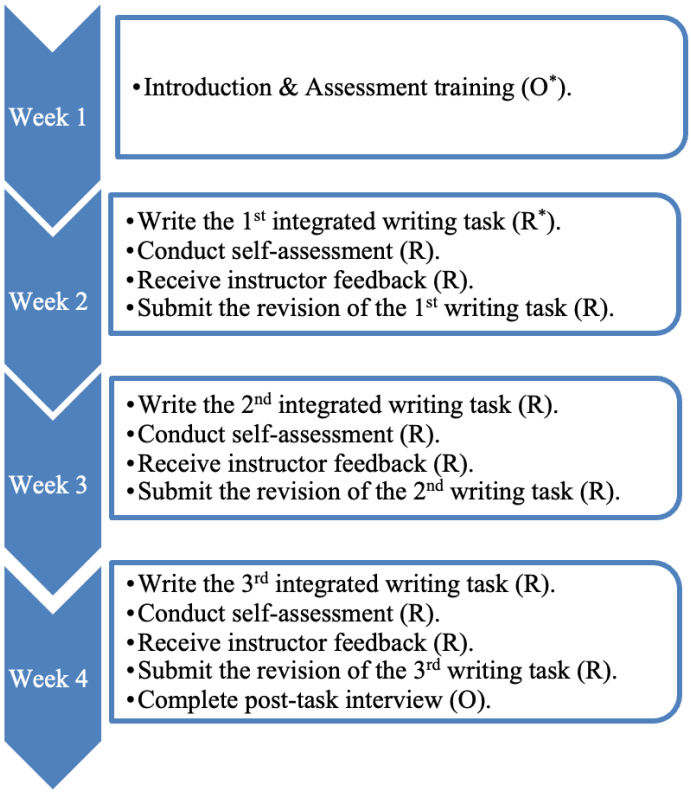



**Figure 3.** *Study schedule.*
**Note*. O: Participants and I met online; R: Participants completed the task remotely.

Before turning to data analysis, it is important to describe the pilot study that preceded the main investigation. The pilot served to trial the research instruments and procedures, identify potential challenges, and inform refinements to the final study design.

### *Pilot study*

A pilot study was conducted with two high-intermediate level ESL undergraduate students enrolled in an academic writing class at a large state university in the United States. As in the main study, pilot participants first joined an introduction and assessment training session. Then,

they wrote the integrated writing essay, conducted self-assessment, responded to my feedback, and submitted a revision. The process writing was repeated two times. A post-task interview was conducted at the end of the study.

The pilot study revealed that high-intermediate level ESL learners did not gain significant benefits from the study, as they were already proficient in incorporating source texts into their writing. Specifically, both participants frequently and successfully employed indirect source use, which aligned with findings from previous research (Gebril & Plakans, 2009; Plakans, 2009a). One contributing factor could be that the participants had already taken the TOEFL iBT or IELTS before entering the ESL program, meaning they had prior exposure to skills integration tasks and had learned to incorporate source texts while preparing for these standardized tests and attending academic writing classes. Given this, the focus for this group of writers should shift to word or sentence-level instruction, with feedback on accuracy, complexity, and variety. Consequently, I revised my target population for the research study to include intermediate or low-intermediate level EFL writers who had never studied in an English-speaking country and were preparing for a language proficiency test.

The purpose of conducting the pilot study was also to collect feedback from participants on the instruments and the study design. Plakans and Gebril (2015) identified the importance of trying out and piloting tasks such as guiding task development, helping with revision, and offering evidence of authenticity. The writing tasks used in this pilot phase were selected from the TOEFL iBT test preparation materials developed by The Princeton Review and Barron's Educational Series. Although the tasks themselves simulate TOEFL iBT integrated writing tasks, the pilot study investigated issues such as the organization, difficulty level, language, length, and topic familiarity of the tasks, as well as clarity of test directions. Three out of four writing tasks were chosen for the research study.

In addition to the writing tasks, participants were first asked to reflect upon the quality of the introductory meeting during the post-task interview. I focused on issues such as clarity, mode, length, and speed of the presentation. They were also asked about their opinions on the assessment checklist such as the organization, validity, and language of the checklist. I also collected participants' opinions on the clarity of the instructions. Opinions and recommendations about the checklist were also sought out from the writing center tutor. Finally, during the pilot study I hoped to try out the interview questions.

In addition to seeking feedback from participants, I kept a research journal to jot down my thoughts throughout the pilot study. These thoughts later helped me refine my instruments and study design. Revisions of the study design and instruments were made before the research study data collection. For a detailed description of the purposes of the pilot study, please see Appendix F.

Based on my reflections in the research journal and feedback from participants in the pilot study, I reviewed and revised all research instruments to enhance the study's quality. Additionally, the pilot study helped improve the introductory meeting by clarifying the presentation of source use and the assessment training for the participants. Revisions to all materials were completed before the research study began. Appendix G provides a summary of my reflections and the participants' feedback from the pilot study.

## Data Analysis

Since all interviews were conducted in Chinese, I first translated the interview scripts into English before transcribing the data. A native Chinese speaker proficient in English was invited to review 20% of the transcription and translation for accuracy. Once he confirmed that the transcription and translations were accurate, I proceeded to complete the remaining dataset.

All data were segmented into idea units, which were defined as the smallest group of words that convey a single meaning (Plakans *et al*., 2019). A colleague familiar with L2 writing research was invited to 20% of the data to ensure the dependability of the coding process. The

agreement rate exceeded 95%. Following this, I coded the remaining data. Then, each idea unit was assigned one or more In Vivo codes, which "prioritize and honor the participant's voice" (Saldaña, 2015, p. 74) to capture key information. The same process was repeated to check inter-rater reliability before I independently coded the remaining data. A summary memo was written for each participant's data to capture key insights after coding their responses. (Saldaña, 2015). I paired the codes and memos for each participant and reviewed them to identify emerging categories and themes.

In addition to analyzing the interview data, I also calculated the agreement between participants' self-assessments and instructor assessments. This calculation assessed how accurately participants evaluated their own work. Triangulating the two sources helped avoid the false assumption that high agreement between self- and instructor assessments necessarily indicated a clear understanding of the assessment criteria. This is important because some participants may have simply checked all the boxes on the checklist without being able to identify the specific instances in their writing that matched the criteria.

## RESULTS

This study forms part of a larger research project investigating whether a learning-oriented assessment approach effectively supports participants' development of source use skills. To address the research questions, I conducted a comprehensive analysis of all participants' data using a conceptually clustered matrix. This practical tool aids in qualitative analysis by organizing rows and columns to group together key themes, variables, concepts, and subtopics for easy reference and summary documentation (Miles *et al*., 2014, p. 173). For this study, I arranged the matrix with all research questions across the top and each participant's performance or perceptions down the left-hand column. This structure enables a quick overview of each participant's profile across the research questions. By reading across the rows, I was able to examine the relationships between responses to each research question. Reading down the columns, I compared subtopics, variables, concepts, and themes within each participant's data as well as across different groups. Findings showed that participants benefited unevenly from the LOA approach (as shown in Figure 4).

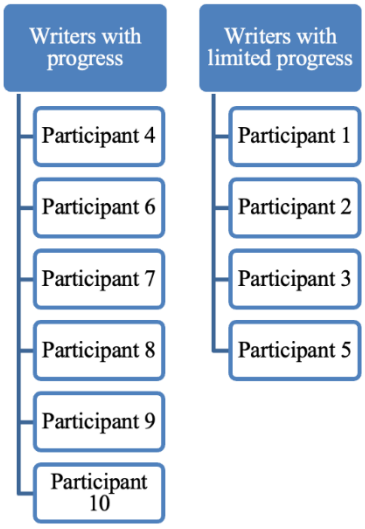



**Figure 4.** *Categories of participants' performance.*

Participants' performances fell into two groups: those who showed progress and those with little progress. Writers who demonstrated progress were able to incorporate source texts implicitly in their writing without any issues of plagiarism by the end of the workshop. In contrast, writers with limited progress continued to overuse direct quotations throughout all writing sessions.

### Writers with Progress

The analysis of the agreement between participants' self-assessment and instructor-assessment revealed that not all writers with progress showed a high level of agreement between their self-assessment and instructor-assessment. Half of these writers demonstrated a lower level of agreement on each writing task. The average agreement among writers with progress was 54.9% (see Table 3). Interview data further explored their perceptions of self-assessment, which were categorized into two main themes: negative experiences and positive experiences. A closer examination of the negative experiences revealed two sub-categories: confusion about evaluation criteria and perceptions of self-assessment as unhelpful.

**Table 3.** *Agreement between Self- and Instructor-assessment.*

| | Task 1 | Task 2 | Task 3 | Average |
|---|---|---|---|---|
| Participant 4 | 25% | 62.5% | 25% | 37.5% |
| Participant 6 | 37.5% | 25% | 50% | 37.5% |
| Participant 7 | 87.5% | 75% | 100% | 87.5% |
| Participant 8 | 62.5% | 87.5% | 75% | 75% |
| Participant 9 | 50% | 62.5% | 50% | 54.2% |
| Participant 10 | 50% | 12.5% | 50% | 37.5% |
| Average | 52.1% | 54.2% | 58.3% | 54.9% |

#### *Negative experience with self-assessment*

**Confusion about Evaluation Criteria.** Although this group of writers showed progress in their ability to use sources throughout the writing workshop, the majority reported confusion regarding the evaluation criteria. For instance, participant 4 demonstrated some incongruences between her self-assessment and my assessment, so I inquired about her understanding of the checklist concepts. She explained that she did not fully grasp the source-use approaches at the start, which led to confusion during self-assessment. Specifically, she believed that direct quotations meant simply replacing a few words from the original text. Similarly, participant 10, who struggled to accurately assess her own performance like participant 4, mentioned that filling out the checklist was complicated and that she had never encountered such an assessment method before.

Interestingly, participant 8, despite showing a higher level of agreement between his self-assessment and the instructor's assessment, also experienced similar confusion. For instance, he was unsure about what constituted a sentence structure different from the original text, leading to inaccuracies in his self-assessment for the first essay. Due to the confusion he faced during the self-assessment process, he recommended adding a section to the checklist where assessors can indicate whether they fully understand the descriptors.

> ***Participant 8:*** "**表格的**话，我觉得是可以接受的。然后就是方框上不是还有一个可以自己写或者自己不写的那个，有一个大框的内容在里面嘛，我觉得那个地方就是可以写一下，自己到底清不清楚自己有没有抄袭，我觉得这样的话对老师啊，或者对阅读我这个那个自我检查报告的这个人可能对我会有一定的了解。"
>
> "About the checklist, I think it is acceptable. And in every section there is a place where you can either write or leave it blank, there is a section in (the checklist), I think that section can ask (students) to write, if they are clear about whether or not they plagiarize in writing, I think in this way, it is better for the teacher, or for the person who reads my self-assessment report, has more understanding about me."

**Perceptions of Self-Assessment as Unhelpful.** In addition to the participants who were unclear about the evaluation criteria, a few others reported not benefiting from the self-assessment process. Participant 6, whose self-assessment showed a low agreement with the instructor's assessment, and participant 7, whose self-assessment showed a high agreement, both expressed that they found the assessment process unhelpful or unnecessary.

> ***Participant 7:*** "其实自我评估这个表我感觉我每次写都是一样的，感觉我，其实，好像没怎末用这个表，每次都是一样的，就是填了一下然后，感觉，没有太大的用处（笑)。
>
> "Actually, I feel that I filled out the same content in the self-assessment checklist every time when using it. I feel, actually, didn't use this checklist too much, the assessment was the same every time, just filled it out, feel, not very useful (laugh)."

Similar to Participant 7, who considered self-assessment unnecessary, Participant 6 felt that self-assessment did not help her address the concerns she identified when evaluating her essays. Specifically, she mentioned that the writing criteria on the checklist, which she found difficult to meet during self-assessment, remained unaddressed. The following excerpt is from Participant 6's interview.

> ***Participant 6:*** "**我**觉得自我评估的表格，其实没啥用，因为（笑）我之前把文章写完后，其实自己写了啥都知道，再去评估一下感觉多了这个步骤。**其**实会的都会，不会的还是不会，就是感觉有点多余。"
>
> "I think the self-assessment checklist, actually isn't useful, because (laugh) after I wrote my essays, I knew what I wrote, so assessing (essays) seemed quite unnecessary. Those I understood were still clear to me, while those I didn't understand were still unclear to me, so I feel (the checklist) was a little unnecessary."

### *Positive experience with self-assessment*

Although many of the writers had negative experiences with self-assessment, a few participants reported benefiting from this innovative approach. For instance, despite Participant 7 finding the checklist useless, she intentionally considered whether she had paraphrased source texts while writing and completed the checklist once her writing was finished. The following excerpt illustrates Participant 7's perspective on the usefulness of self-assessment.

> ***Participant 7:*** "…**就是写作的**时候刻意地想一下有没有paraphrasing，**然后之后填一下那个表格**…"
>
> "…Just when writing, I would purposefully think about whether I paraphrased, then filled out the checklist..."

Participant 4 said that self-assessment offered her another opportunity to reread her essays.

> ***Participant 4:*** "**写完之后，我重新**读一遍我的文章，并且，让我再读一遍文章。"
>
> "After writing, I read my essay for another time, and, let me read the essay one more time."

To summarize, although most writers showed progress in their source use ability throughout the writing workshop, the self-assessment process did not significantly contribute to their deeper understanding of source use or to the development of their learner autonomy and confidence. In fact, many participants found self-assessment to be confusing and unhelpful.

## Writers With Limited Progress

Throughout the writing workshop, some participants showed minimal progress. Specifically, participants 1, 2, 3, and 5 showed little improvement. The essays written by participants 1 and 5 were characterized by a heavy reliance on explicit source use across all three writing sessions.

In contrast, participants 2 and 3 exhibited very few instances of textual borrowing in their essays, categorizing them as writers with limited progress.

Triangulating multiple data sources revealed several reasons for participants' limited progress. For instance, the progress of participants 1 and 5 was significantly hindered by their lack of understanding of the assessment criteria, which led to confusion throughout the three writing sessions. In contrast, participants 2 and 3 (whose English proficiency ranged from low-intermediate to high-intermediate) demonstrated excellent source use in almost every draft, leaving little opportunity for further improvement in this area.

The following sections will provide a detailed analysis of participants 1 and 5. The two highly proficient writers were excluded from detailed source-use analysis because self-assessment influenced them more in other areas, such as language accuracy and complexity. This conclusion was drawn based on an examination of their essays and their feedback during the interviews.

### *Negative experience with Self-assessment*

**Confusion about Evaluation Criteria.** Similar with writers with progress, both participant 1 and participant 5 reported that they were very confused about the assessment checklist. Participant 1 felt that sometimes he had met the criteria, so he indicated this on his assessment, while sometimes he did not feel that he had accomplished the criteria on the assessment checklist. Participant 5 also reported a similar feeling in the following excerpt:

***Participant 5:*** "…**就是主要是**刚开始的话这个评估表格我实在是没有看懂…"

"…So, I really couldn't understand the assessment checklist at the beginning, …"

An examination of the agreement between self-assessment and instructor assessment also yielded the same result (see Table 4). For instance, in participant 1's self-assessment checklist, he checked the box "When I directly quoted materials from the readings, I properly referred to and cited the author(s)" but marked an "X" in all other boxes.

**Table 4.** *Agreement between self-and instructor-assessment.*

| | Task 1 | Task 2 | Task 3 | Average |
|---|---|---|---|---|
| Participant 1 | 25% | 0% | 0% | 8.3% |
| Participant 5 | 62.5% | 37.5% | 12.5% | 37.5% |
| Average | 43.8% | 18.8% | 6.3% | 22.9% |

However, he had neither paraphrased nor summarized the source texts in his essay. Most of the information presented was not sourced from the reading passages.

## DISCUSSION and CONCLUSION

Research in both general education and language testing has consistently shown that self-assessment can support learning, but only when learners are provided with sufficient training to understand assessment criteria (Ross, 1998, 2006). This study's findings echo those results (as shown in Table 5). Writers who demonstrated progress also showed higher agreement between their self-assessment and instructor ratings, suggesting that they had internalized the criteria more effectively. By contrast, writers with limited progress displayed weaker agreement, indicating confusion about the criteria and reinforcing Carless's (2007) and Jones and Saville's (2016) arguments that transparency of expectations is essential.

The interview data revealed similar findings. Although most learners showed some progress, many, particularly those with limited progress, reported confusion about the checklist descriptors, which prevented them from accurately assessing their work or fully understanding feedback. Two participants from the group of writers with progress considered self-assessment

unhelpful or unnecessary. One participant said that although self-assessment helped her identify the issues that her essays failed to meet, it cannot provide her with solutions to address those issues, thus not serving to help her improve her writing before she handed in her essays. Among the two writers who thought self-assessment beneficial, one participant made the same self-assessment every time, although she would purposefully monitor herself during the writing stage. Another participant who made progress throughout the study said that self-assessment offered her another opportunity to reread her essays.

**Table 5.** *Average Agreement between Self- and Instructor-assessment.*

| | Task 1 | Task 2 | Task 3 | Average |
|---|---|---|---|---|
| Writers with progress (6) | 52.1% | 54.2% | 58.3% | 54.9% |
| Writers with limited progress (2)* | 43.8% | 18.8% | 6.3% | 22.9% |

*Note**: The numbers in the parentheses represent the number of writers with in-depth analyses in each group.

Regarding the group of writers for whom self-assessment did not work, all participants reported being confused about the assessment criteria. Among those participants, two came from the group of writers with limited progress. This finding corroborates findings that learners must clearly understand the criteria by which their work will be judged (Carless, 2007; Jones & Saville, 2016). Frederiksen and Collins (1989) also emphasized the importance of transparency in supporting learners to effectively evaluate their work on their own. Carless (2007) argued that learners should be aware of the evaluation criteria and should be monitoring their progress towards these standards. In order for learners to have such an awareness and behavior, they have to be equipped with solid understandings and experiences of assessment. Otherwise, there is a lack of transparency among learners about what counts as distinct features of outstanding performance (Frederiksen & Collins, 1989).

Furthermore, Rolheiser's (1996) model of self-assessment clearly indicates the prerequisites for learners to demonstrate achievement, which is a combination of goals and effort. Because many of the participants in this study were not equipped with a clear understanding of their learning goals, their effort alone could not help them make learning gains, thus failing to accurately evaluate their work, develop self-confidence, and further set higher goals and commit more effort to subsequent learning.

As discussed in the literature review, Rolheiser and Ross (2011) developed a four-stage model for teaching learners how to self-assess their work. The assessment training I designed closely followed this model (which I discovered after data collection), but I did not allocate enough time to fully implement each stage, particularly the third step. This was partly because participants were involved in the study for research purposes rather than as my own students, which would have given me more opportunities to provide instruction and practice on self-assessment before they actually applied it.

The study findings suggest viewing the understanding and experiences of self-assessment that teachers and learners have as having a linear relationship with the teachers' understanding and experiences as influencing learners' understanding and experiences. In fact, teachers' understandings of learners are vital for the success of getting learners involved in self-assessment. Because of my role as a workshop tutor instead of their classroom instructor, I did not have a solid understanding of my participants despite the efforts I had made to get to know them at the instigation of data collection through communication with the school principal and the information provided on the school website. My limited understanding of my participants negatively affected the quality of the training that I provided to writers with the goal of preparing them to be capable self-assessors, decreasing the effectiveness of their involvement in self-assessment.

The study findings also shed light on the validity of self-assessment. Previous scholarship shows that variables that can affect the accuracy of self-assessment include learners'

characteristics, such as learners' second or foreign language proficiency levels, with higher-proficiency learners assessing their performance more accurately than lower-proficiency learners (Brantmeier *et al*., 2012). In this study, the two writers who made limited progress had the least level of English proficiency, which may explain their lack of accuracy in self-assessment. The way in which learners comprehend and respond to the evaluation criteria is also associated with their performance on self-assessment (Heilenman, 1990). The agreement check between participant- and instructor-assessment revealed that writers with progress had a higher agreement with instructor-assessment than writers with limited progress. This may indicate that writers with progress had a clearer understanding of the criteria listed on the assessment checklist, thus performing better on self-assessment, and using the assessment tool to monitor performance and enhance learning. On the contrary, confusion about the evaluation criteria resulted in a lack of accuracy in self-assessment for writers with limited progress, which inhibited them from using assessment to promote learning.

Another factor that positively associates with accuracy in self-assessment is using learners' first language to develop the assessment tool (Oscarson, 1997). In this study, the assessment checklist was written in English since I employed plain language to write the evaluation criteria and explained those criteria in Chinese during the assessment training to increase understanding. However, participants would have had a better understanding of the assessment checklist and utilized the tool to regulate their performance if it was written in their first language or in a combination of English and their first language. The validity of self-assessment would have been increased through carefully designed learner training and the assessment tool (Ross, 2006). In terms of practicality of self-assessment, it also suffered given my role as a temporary instructor and the online mode of holding the entire writing workshop which made communication and interaction between participants and me difficult. Future studies should take learners' first language into consideration when designing the assessment tool and ensure that classroom teachers are involved as part of the study.

Although some study limitations have been mentioned, it is still worth discussing them further. First, participants in this study were international high school students in China preparing for overseas undergraduate studies and English proficiency tests (TOEFL or IELTS), enrolled in test-preparation courses and familiar with reading-to-write tasks. Despite prior exposure, they were still relatively new to applied source use, which limited the effectiveness of the LOA approach, as self-assessment preceded formal instruction. The virtual classroom format further constrained LOA's impact by eliminating in-person interactions that support learning processes, forcing the study to prioritize essay products over interactive learning experiences. Additionally, participants completed only three writing tasks, which may have been insufficient to capture meaningful development in source use and their potential growth in self-assessment skills. Future research should involve students receiving formal source-use instruction, incorporate face-to-face interactions when feasible, and adopt longitudinal designs to allow more practice and richer observation of self-assessment development. Additional, the study underscores the importance of investigating how self-assessment can be tailored to different proficiency levels and how bilingual or first-language support might enhance its effectiveness.

In conclusion, this study explored whether a group of secondary school EFL writers were able to enhance their source use ability by participating in self-assessment in a one-month writing workshop. Findings show that self-assessment failed to build participants' evaluative expertise or foster self-regulated learning, largely due to their lack of clear understanding of the evaluation criteria. Future instructors should devote sufficient time to training learners—especially beginners in self-assessment—using models such as Rolheiser and Ross (2011) to maximize learning outcomes. Future research should involve trained students, incorporate face-to-face interactions, use longitudinal designs, and examine tailoring self-assessment to proficiency levels and bilingual support.

## Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author.

## Orcid

Fang Wang https://orcid.org/0000-0002-7164-9814

## REFERENCES

AlFallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer-assessment. *System, 32*(3), 407-425. https://doi.org/10.1016/j.system.2004.04.006

Alverno College Faculty. (2000). *Self assessment at Alverno College*. Alverno College Institute.

Andrade, M.S., & Evans, N.W. (2012). *Principles and practices for response in second language writing: Developing self-regulated learners*. Routledge.

Bachman, L.F., & Palmer, A.S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6*(1), 14-29. https://doi.org/10.1177/026553228900600104

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347-364. https://doi.org/10.1007/BF00138871

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151-167. https://doi.org/10.1080/713695728

Boud, D. (2003). *Enhancing learning through self-assessment*. Routledge/Falmer. https://doi.org/10.1080/713695728

Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me?: Individual self-assessment by skill and level of language instruction. *System, 40*(1), 144-160. https://doi.org/10.1016/j.system.2012.01.003

Butler, Y.G., & Lee, J. (2006). On-task versus off-task self-assessments among Korean elementary school students studying English. *The Modern Language Journal, 90*(4), 506-518. https://doi.org/10.1111/j.1540-4781.2006.00463.x

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in education and teaching international, 44*(1), 57-66. https://doi.org/10.1080/14703290601081332

Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education, 69*, 963-976. https://doi.org/10.1007/s10734-014-9816-z

Carless, D., Joughin, G., & Mok, M. (2006a). Learning-oriented assessment: Principles and practice. *Assessment and evaluation in Higher Education, 31*(4), 395-398. https://doi.org/10.1080/02602930600679043

Carless, D., Joughin, G., & Liu, N.F. (2006b). *How assessment supports learning: Learning-oriented assessment in action* (Vol. 1). Hong Kong University Press.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing, 10*(1), 5-43. https://doi.org/10.1016/j.asw.2005.02.001

Cumming, A., Lai, C., & Cho, K. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of Second Language Writing, 31*, 1-17. https://doi.org/10.1016/j.jslw.2016.06.001

De Saint-Leger, D. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals, 42*, 158-178. https://doi.org/10.1111/j.1944-9720.2009.01013.x

De Saint-Leger, D., & Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System, 37*(2), 269-285. https://doi.org/10.1016/j.system.2009.01.001

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3), 331-350. https://doi.org/10.1080/03075079912331379935

Falchikov, N. (2006). *Improving assessment through student involvement: Practical solutions for learning in higher and further education*. Routledge/Falmer.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27-32. https://doi.org/10.3102/0013189X018009027

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*(1), 47-84.

Gebril, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: Lexical diversity, textual borrowing and proficiency. *Journal of English for academic purposes, 24*, 78-88. https://doi.org/10.1016/j.jeap.2016.10.001

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hauser, G.M., Koutouzos, D.W., & Olson, G.E. (2005). The standards-based school leader digital portfolios. *International Journal of Learning*, *12*(7), 311-316.

Heilenman, L.K. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing, 7*(2), 174-201. https://doi.org/10.1177/026553229000700204

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

Jamieson, J.M., Eignor, D., Grabe, W., & Kunnan, A.J. (2008). Frameworks for a New TOEFL. In C. Chapelle, M. Enright, & J. Jamieson (Eds.), *Building a validity argument for the Test of English as a Foreign Language* (pp. 55-96). Routledge.

Jones, N., & Saville, N. (2016). *Learning oriented assessment: A systemic approach*. Cambridge University Press..

Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, *15*(4), 261-278. https://doi.org/10.1016/j.jslw.2006.09.006

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

Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *The Canadian Modern Language Review*, *64*(1), 199-213. https://doi.org/10.3138/cmlr.64.1.199

Liu, X., & Li, L. (2014). Assessment training effects on student assessment skills and task performance in a technology-facilitated peer assessment. *Assessment & Evaluation in Higher Education*, *39*(3), 275-292. https://doi.org/10.1080/02602938.2013.823540

Lowry, C. (2016). *Choosing & using sources: A guide to academic research.* Ohio State University.

MacIntyre, P.D., Noels, K.A., & Clement, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning*, *47*(2), 265-287. https://doi.org/10.1111/0023-8333.81997008

Miles, M.B., Huberman, A.M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Min, H.T. (2005). Training students to become successful peer reviewers. *System, 33*(2), 293-308. https://doi.org/10.1016/j.system.2004.11.003

Office of the Vice President for Research. (2020, March). *Temporary policy on human subjects research-related visits during COVID-19 outbreak effective March 16*. The University of Iowa. https://research.uiowa.edu/news/2020/03/temporary-policy-human-subjects-research-related-visits-during-covid-19-outbreak

Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham & D. Corson (Eds.), *The encyclopedia of language and education. Vol. 7: Language testing and assessment* (pp. 175-87). Kluwer.

Oscarson, M. (2013). Self-assessment in the classroom. In A. J. Kunnan (Ed.), *The companion to language assessment: Approaches and development* (pp. 712-729). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118411360.wbcla046

Oskarsson, M. (1978). *Approaches to self-assessment in foreign language learning*. Council of Europe, Council for Cultural Cooperation.

Patri, M. (2002). The influence of peer feedback on self-assessment and peer assessment of oral skills. *Language Testing, 19*(2), 109-131. https://doi.org/10.1191/0265532202lt224oa

Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing, 13*(2), 111-129. https://doi.org/10.1016/j.asw.2008.07.001

Plakans, L. (2009a). Discourse synthesis in integrated second language writing assessment. *Language Testing, 26*(4), 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*(4), 252-266. https://doi.org/10.1016/j.jeap.2009.09.002

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., & Gebril, A. (2015). *Assessment myths: Applying second language research to classroom teaching*. University of Michigan Press. https://doi.org/10.3998/mpub.5056216

Plakans, L., Liao, J.T., & Wang, F. (2019). "I should summarize this whole paragraph": Shared processes of reading and writing in iterative integrated assessment tasks. *Assessing Writing, 40*, 14-26. https://doi.org/10.1016/j.asw.2019.03.003

Purpura, J.E. (2016). Second and foreign language assessment. *Modern Language Journal, 100*(Supplement), 190-208. https://doi.org/10.1111/modl.12308

Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research, 17*(1), 67-89. https://doi.org/10.1177/1362168812459151

Rolheiser, C. (Ed.). (1996). *Self-evaluation: Helping students get better at it*. Visutronx.

Rolheiser, C., & Ross, J.A. (2001). Student self-evaluation: What research says and what practice shows. In R.D. Small & A.P. Thomas (Eds.), *PlainTalk about education: Transforming crisis into success* (pp. 43-57). Center for Development & Learning.

Ross, J.A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation, 11*(1), Article 10. https://doi.org/10.7275/9wph-vv65

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1-20. https://doi.org/10.1177/026553229801500101

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage Publication.

Sadler, R. (2002) Ah! … So that's 'quality', In P. Schwartz & G. Webb (Eds), *Assessment: case studies, experience and practice from higher education*. Kogan Page.

Schuessler, J.N. (2010). *Self assessment as learning: Finding the motivations and barriers for adopting the learning-oriented instructional design of student self assessment* [Unpublished doctoral dissertation]. Capella University.

Shi, L. (2010). Plagiarism in second-language writing. *Language Teaching, 43*(3), 269-296. https://doi.org/10.1017/S0261444814000056

Suzuki, Y. (2015). Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition. *Language Testing, 32*(1), 63-81. https://doi.org/10.1177/0265532214541885

Weigle, S.C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing, 9*(1), 27-55. https://doi.org/10.1016/j.asw.2004.01.002

Weir, C.J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan. https://doi.org/10.1057/9780230514577

Wright, W.E. (2019). *Foundations for teaching English language learners: Research, theory, policy, and practice* (3rd ed.). Caslon Publishing.

Zhang, P., Fan, J., & Jia, W. (2022). Assessing L2 integrated writing self-efficacy: Development and validation of a scale. *Assessing Writing, 54*, Article 100665. https://doi.org/10.1016/j.asw.2022.100665

# APPENDICES

**Appendix A**. *A sample of the integrated writing task.*

**Writing Task 1**

Below are two sides about a topic. Read the two passages below. Then, develop a persuasive argument for your position. Make sure to use **at least three points of information** from the reading texts to support your opinion. The essay should be about 250 words. Your writing will be rated based on how effectively you:

- Incorporate reading materials successfully
- Refer to authors accurately

**Some people support the suppression of forest fires, while others cast doubt on it. What is your opinion? Use reasons and/or examples to support your writing.**

**Please read carefully before starting the task:**

- The two passages should give you some general ideas about the topic.
- You can go back to the readings to check information once you start writing.
- When borrowing ideas and/or examples from the text, make sure to include references to the readings.
- You can use any tools (e.g., online dictionary). However, do not seek out any help from other people or plagiarize from other sources. The purpose of this study is to help you improve your academic writing skills.

**Reading 1 written by Benjamin Lee, 2010, Natural Environment, p. 10**

Wilderness management has advanced greatly over the last century, due in part to such practices as the control of forest fires and limitations on cutting trees. Monitoring forests for small brushfires is easier with aircraft, as is the use of large amounts of water and advanced chemical fire extinguishers to prevent fires from spreading.

There are many goals and benefits of decreasing the amount of destruction by fires and cutting. One is simply the longer lives and improved health of trees. In some areas of hickory and oak forest on the Eastern Seaboard, because of fire suppression, trees have evolved to grow very tall so as to reach more sunlight.

There is evidence of the positive effects of fire suppression closer to the ground as well. Vines and low bushes that would be burned out in a forest fire can survive and grow when fires are suppressed. There is a more indirect way fires harm plant life. Scientists conducted chemical tests on areas that have recently suffered from forest fires. They found that burning decreases the overall amount of nutrients in the soil. Suppressing fires prevents such a decrease. Plants can grow without disturbance.

Finally, wildlife can benefit. In the eastern hickory and oak forests, the suppression of fires has meant that forest animals — ranging from small insects and birds to large deer and, bears — are not burned to death. Deer populations, in particular, have increased quite a lot.

**Reading 2 written by Margaret Cumming, 2014, Current Issues on Environment, p. 67**

For years, forest fires were seen as damaging, and forest managers put a lot of effort into preventing them. But it turns out that fire suppression may have harmful long-term effects on the forests it is supposed to protect.

For instance, mature oaks have grown so thickly in some places that little light reaches the forest floor. However, young oak trees need light in order to grow properly. The lack of light has meant that new oaks are not maturing rapidly enough to replace the older oaks. It also means that other tree species that do not need so much light, such as maples, are invading oak and hickory forests and competing for resources.

There are competition problems at ground level too. Forest fires, both natural and artificial, can burn off some of the plants on the forest floor before they could grow into a dense group of trees or bushes. Now they run wild over the ground. Again, that means it is hard for young trees and other native plants to grow.

Then there is the increase in the deer population — this partially results from the lack of forest fires and limitations on hunting. However, the problem is that deer like to eat the leaves off young oak trees. If one of those oak seedlings somehow does manage to survive, despite the shade and all the other plants competing for nutrients, it is likely to be killed by having its leaves eaten.

Finally, scientists are now finding that forest fires release nutrients from the plants and animals that are burned. That means that, even though the total amount of nutrients is decreased, there can actually be more nutrients available on the soil surface for plants that are trying to grow back afterward.

**Appendix B.** *A sample of the essay paper*.

Instructions:

1. Make sure to use **at least three points of information** from the reading texts to support your opinion.
2. The essay should be about **250 words**.
3. Your writing will be rated based on how effectively you incorporate reading materials successfully and refer to authors accurately.

*Some people support the suppression of forest fires, while others cast doubt on it. What is your opinion? Use reasons and/or examples to support your writing.*

[Please type your writing below]

(Word count: )

**Appendix C**. *Assessment checklist.*

| Assessment Checklist |
| --- |
| Below is a brief checklist to help you with self-assessment. The checklist has three sections, but you do not need to complete each section if you did not use a particular integration style.<br>When conducting self-assessment, please check (√) the box if you think your essay meets that criterion and cross (X) the box if you think your essay does not meet that criterion. You can provide some comments for yourself if needed. AA |
| If direct quotation(s) were used in your essay, please check if you met the following criterion:<br>□ I properly referred to and cited the author(s).<br>□ I used direct quotation(s) only when the source's exact language is particularly effective.<br>Comments (optional):<br><br>If paraphrasing was used in your essay, please check if you met the following criteria:<br>□ I properly referred to and cited the author(s).<br>□ I did not maintain the way the original author constructed sentence(s).<br>□ I did not use the author's language for more than three words.<br>Comments (optional):<br><br>If summarizing was used in your essay, please check if you met the following criteria:<br>□ I properly referred to and cited the author(s).<br>□ I did not maintain the way the original author constructed sentence(s).<br>□ I did not use the author's language for more than three words.<br>Comments (optional): |

**Appendix D**. *Interview questions.*

Thank you very much for taking the time to participate in this study. As the study is ending, I would like to ask you a few questions about your personal experience with the study. The purpose of this interview is to gather insights that can offer practical suggestions for teachers, researchers, and those interested in conducting similar studies in the future, particularly those advocating for the adaptation of learner self-assessment to enhance learning. The interview will take about ten minutes, and you are welcome to skip any questions you do not wish to answer.

---

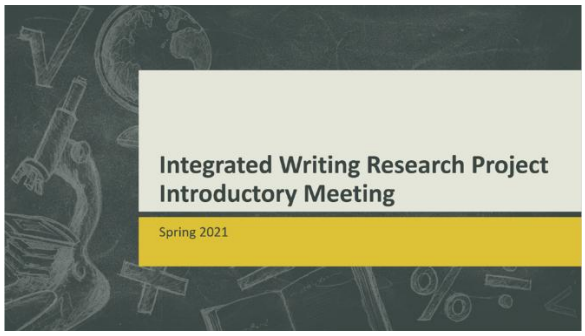
Part I. Basic information

1. What is your grade level?
2. Have you been exposed to integrated writing tasks at school?
3. What have you learned about integrated writing essays from school?
4. What are your most recent English test scores and part scores for reading and writing? (e.g., TOEFL, IELTS)
5. Are you currently enrolled in any test preparation classes?
6. What have you learned about integrated writing essays from the test preparation classes?

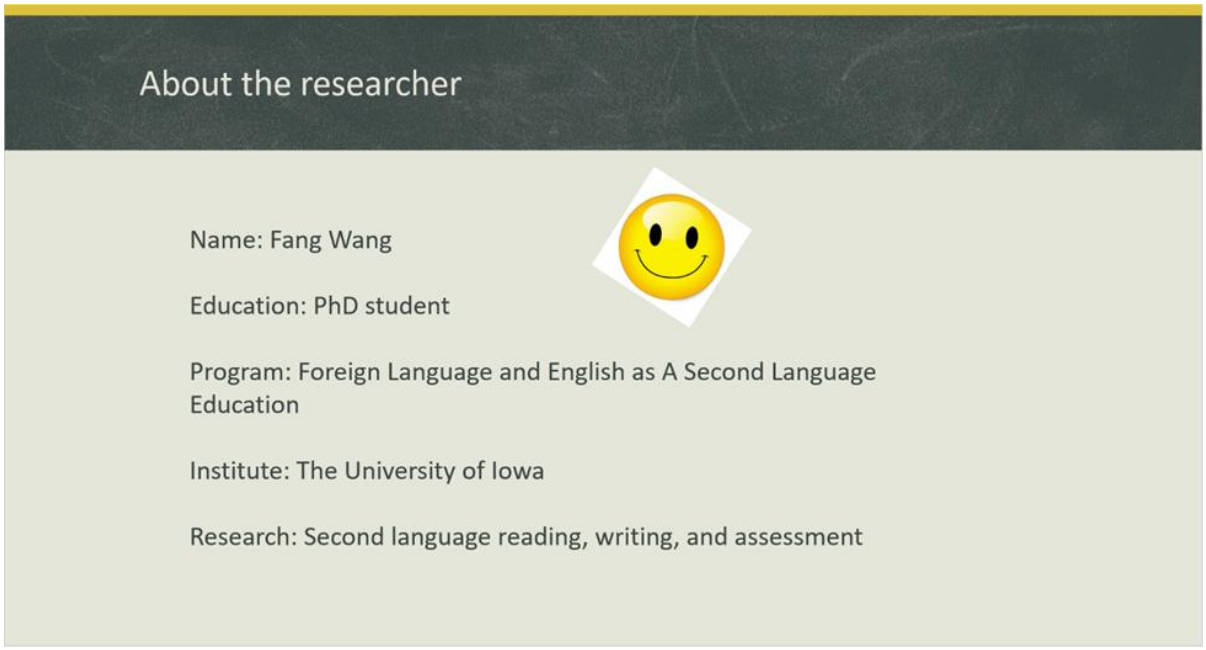
Part II. Questions about the study

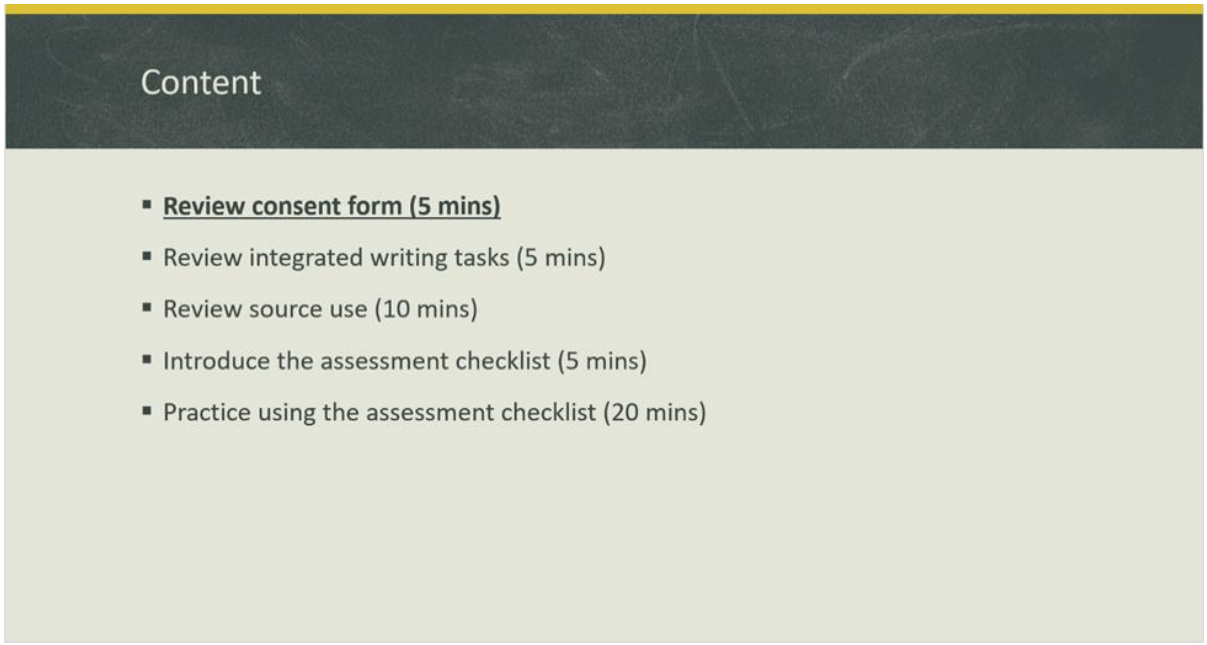
1. Overall, how would you describe your experience with the introductory meeting?
2. What is the most important takeaway you gained from the introductory meeting?
3. Do you have any suggestions for improving the introductory meeting?
4. Overall, how would you describe your experience with self-assessment?
5. Do you have any suggestions about the assessment checklist?
6. Do you have further suggestions for this study?
7. Is there anything that you would like to know?

**Appendix E.** *Powerpoint slides for the introductory meeting.*







## Review consent form

We invite you to participate in a research study. The purpose of the study is to investigate the effectiveness of using learner-oriented assessment on source use in integrated reading-to-write tasks in academic settings. We are recruiting fifteen subjects for a research study (six-hour time commitment).

If you agree to participate, we would like you to attend a brief introductory meeting which will introduce the research study, integrated reading-to-write tasks and source use, as well as the assessment checklist which you will use to self-assess your essays and receive my assessment and feedback. Then, you will write three integrated reading-to-write tasks. Your essays will be evaluated by the researcher and yourself. You will also be asked to turn in a revised draft for each reading-to-write task. You are free to skip any writing activity that you would prefer not to participate in. At the end of the study, you will be invited for a post-task interview which will collect your thoughts regarding the effectiveness of learner-oriented approach on your performance of integrated writing tasks. The study will last for about one month and your participation will require about six hours throughout the study period. The introductory meeting and the post-task interview will take place via Zoom (the latter will be audio recorded, with your permission). All other activities will be completed at your convenience. The interview recordings will be transcribed without your identity, and stored on a secure server; the data will be deleted when this research study is complete.


## Review consent form

We invite you to participate in a research study. The purpose of the study is to investigate the effectiveness of using learner-oriented assessment on source use in integrated reading-to-write tasks in academic settings. We are recruiting fifteen subjects for a research study (six-hour time commitment).

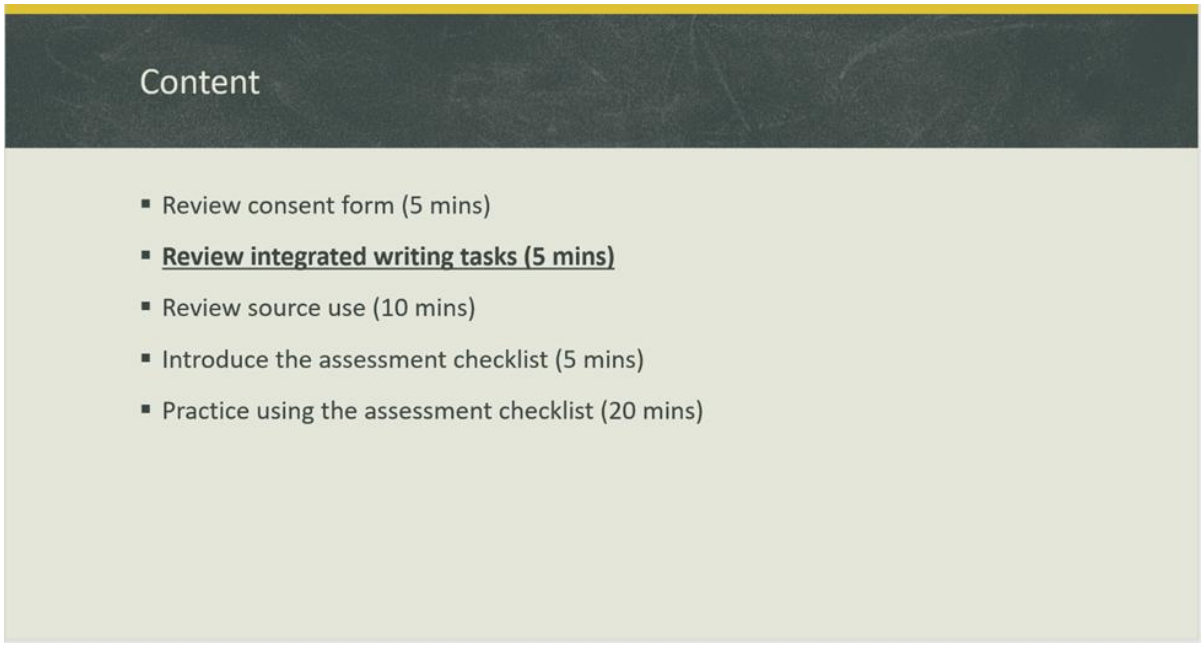
If you agree to participate, we would like you to attend a brief introductory meeting which will introduce the research study, integrated reading-to-write tasks and source use, as well as the assessment checklist which you will use to self-assess your essays and receive my assessment and feedback. Then, you will write three integrated reading-to-write tasks. Your essays will be evaluated by the researcher and yourself. You will also be asked to turn in a revised draft for each reading-to-write task. You are free to skip any writing activity that you would prefer not to participate in. At the end of the study, you will be invited for a post-task interview which will collect your thoughts regarding the effectiveness of learner-oriented approach on your performance of integrated writing tasks. The study will last for about one month and your participation will require about six hours throughout the study period. The introductory meeting and the post-task interview will take place via Zoom (the latter will be audio recorded, with your permission). All other activities will be completed at your convenience. The interview recordings will be transcribed without your identity, and stored on a secure server; the data will be deleted when this research study is complete.

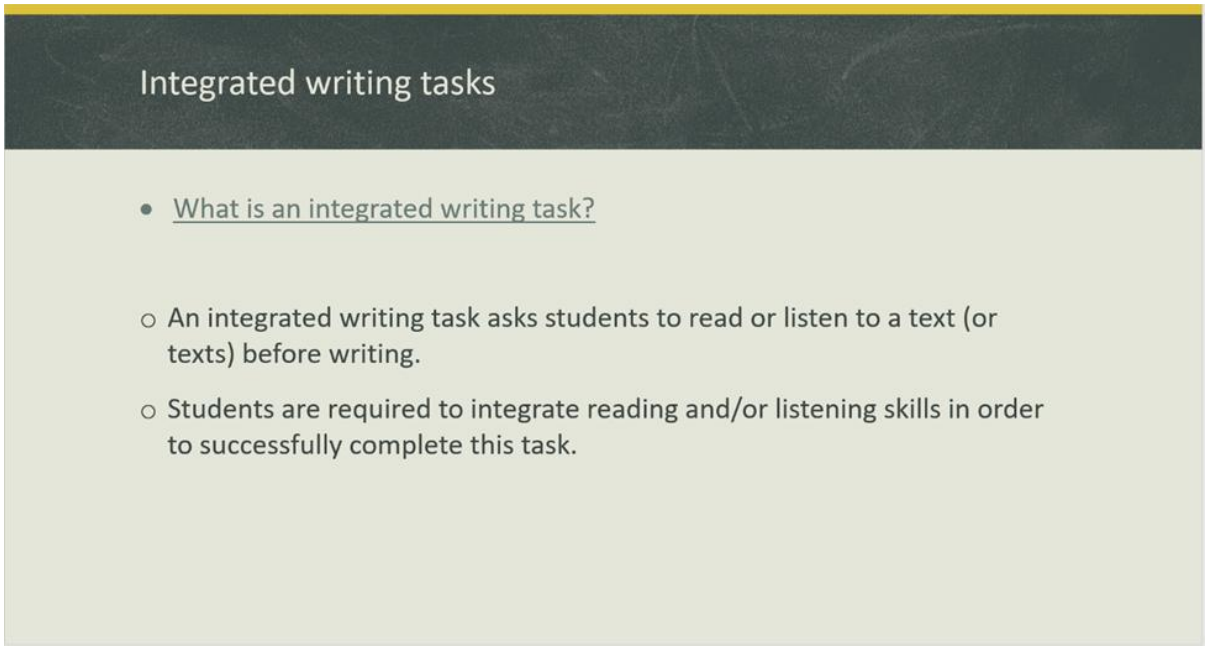
You will be paid for participating in this research study. You may also need to provide your address if you prefer your check to be sent through the mail. You will receive a $150 check for completing all activities in this study at the end of the study. If you withdraw midway through the study, you will receive $75 compensation.

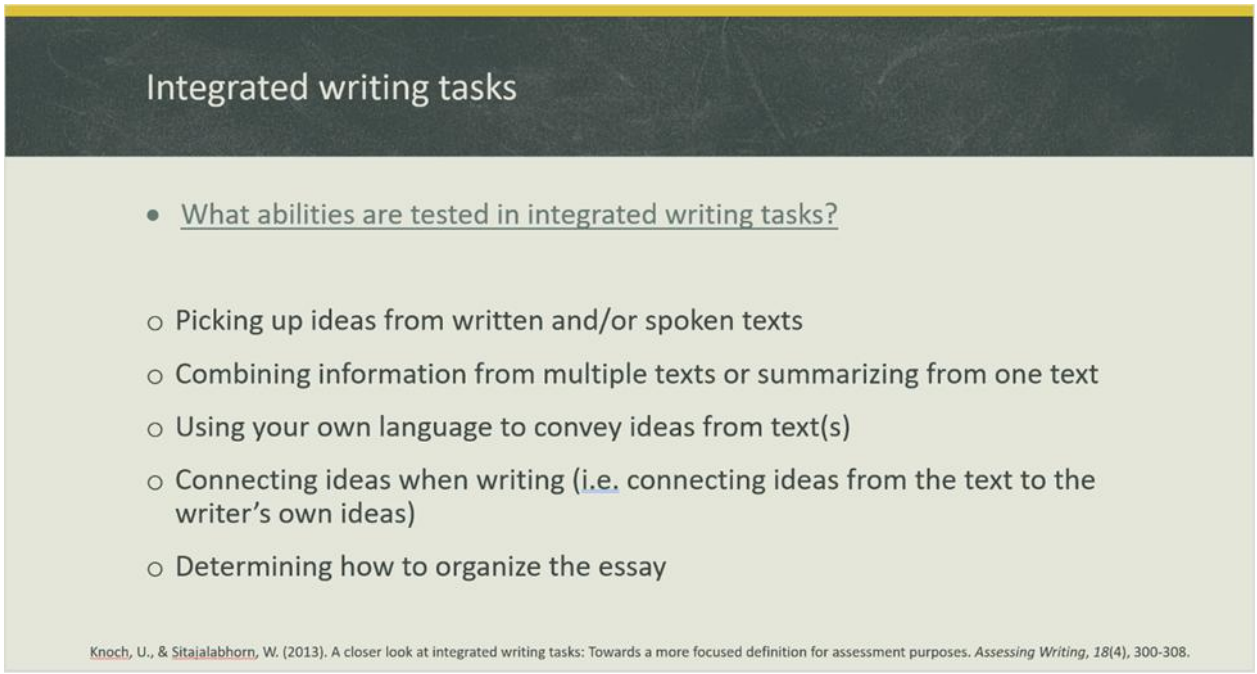
Any information that you provide will remain confidential. A study ID will be used in place of your personal name and your data will not be used for future research. If we write a report about this study, we will do so in such a way that you cannot be identified.

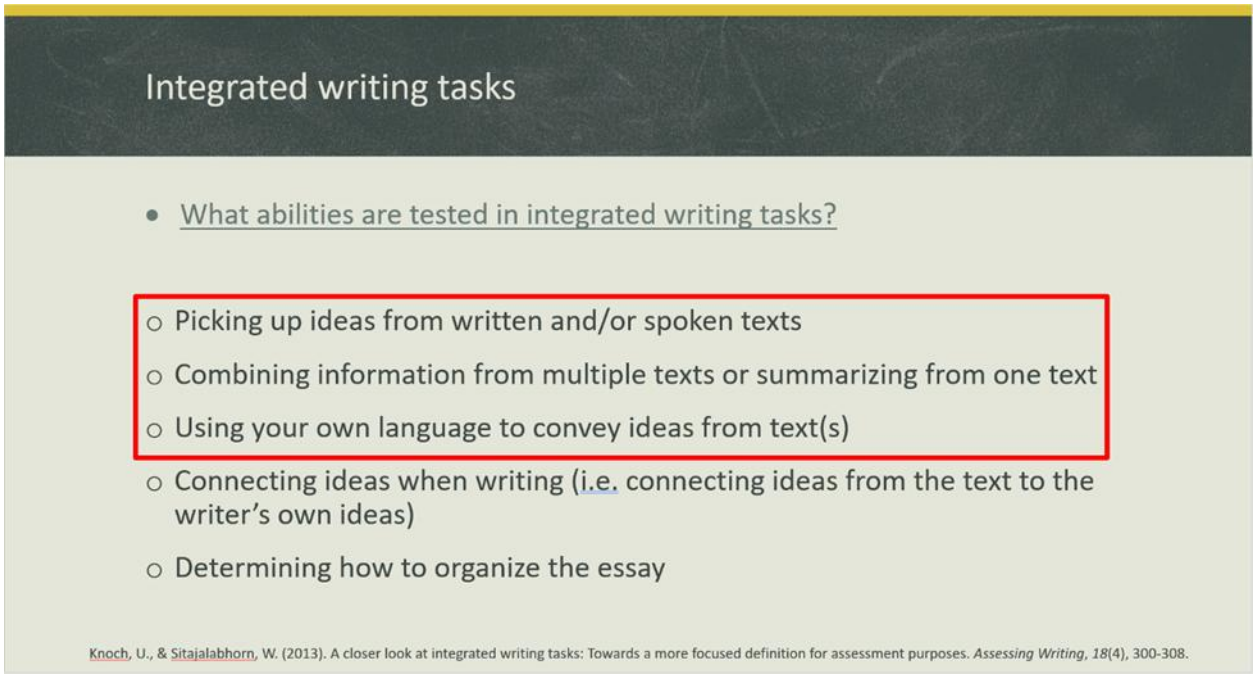
Taking part in this research study is completely voluntary. If you have questions about the rights of research subjects, please contact the Human Subjects Office, 105 Hardin Library for the Health Sciences, 600 Newton Rd, The University of Iowa, Iowa City, IA 52242-1098, (319) 335-6564, or e-mail irb@uiowa.edu.

Thank you very much for your consideration.


Content
Review consent form (5 mins)
Review integrated writing tasks (5 mins)
Review source use (10 mins)
Introduce the assessment checklist (5 mins)
Practice using the assessment checklist (20 mins)


Integrated writing tasks
What is an integrated writing task?
An integrated writing task asks students to read or listen to a text (or texts) before writing.
Students are required to integrate reading and/or listening skills in order to successfully complete this task.

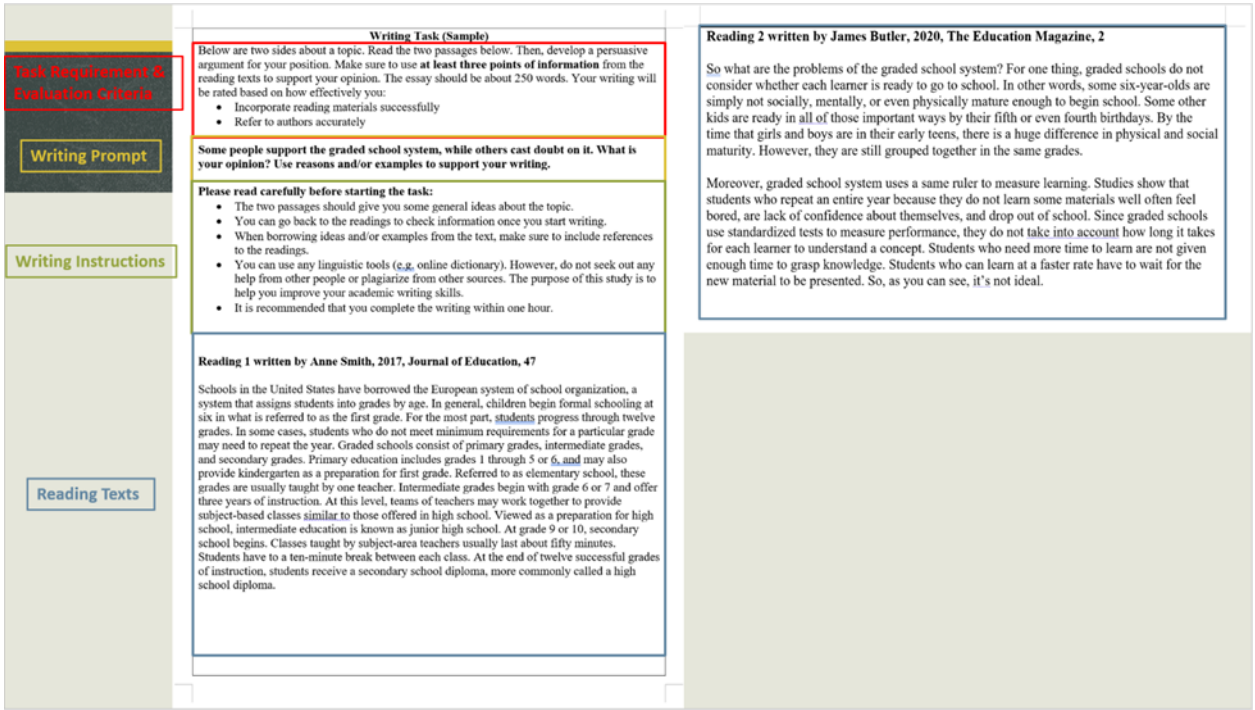

Integrated writing tasks
What abilities are tested in integrated writing tasks?
Picking up ideas from written and/or spoken texts
Combining information from multiple texts or summarizing from one text
Using your own language to convey ideas from text(s)
Connecting ideas when writing (i.e. connecting ideas from the text to the writer's own ideas)
Determining how to organize the essay
Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focused definition for assessment purposes. Assessing Writing, 18(4), 300-308.

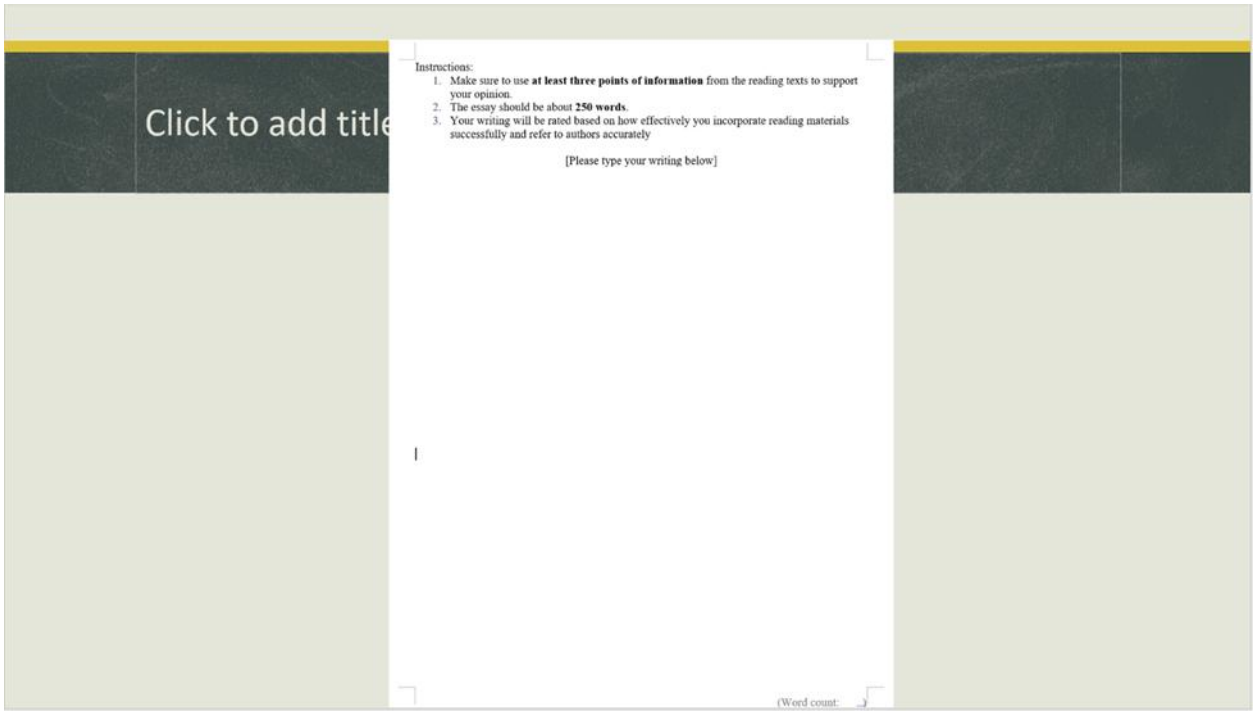

Integrated writing tasks
What abilities are tested in integrated writing tasks?
Picking up ideas from written and/or spoken texts
Combining information from multiple texts or summarizing from one text
Using your own language to convey ideas from text(s)
Connecting ideas when writing (i.e. connecting ideas from the text to the writer's own ideas)
Determining how to organize the essay
Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focused definition for assessment purposes. Assessing Writing, 18(4), 300-308.

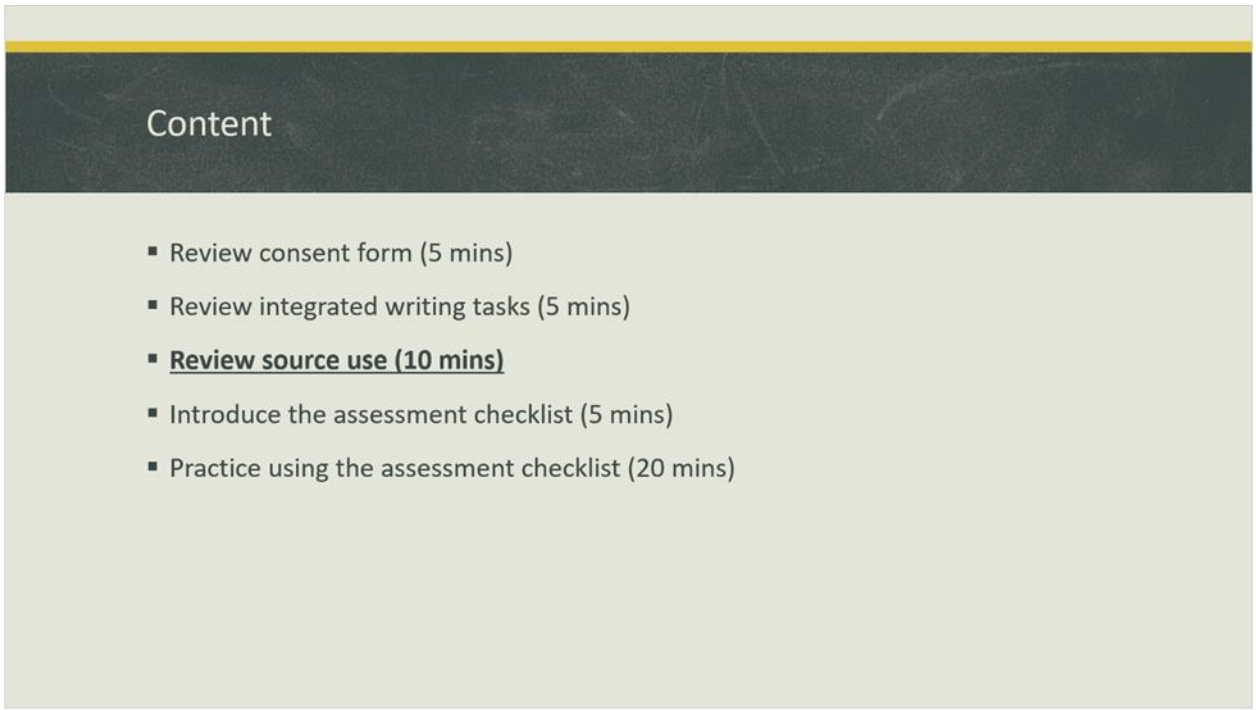
## Integrated writing tasks

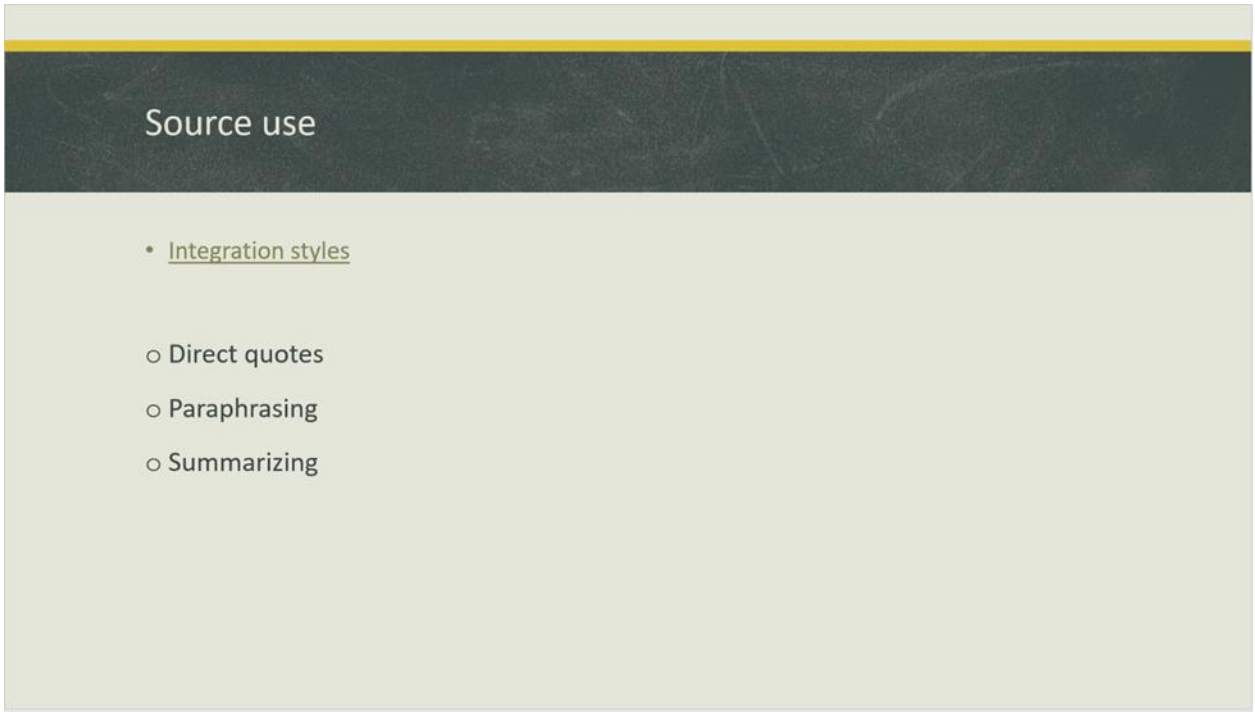
- What are some examples of integrated writing tasks?

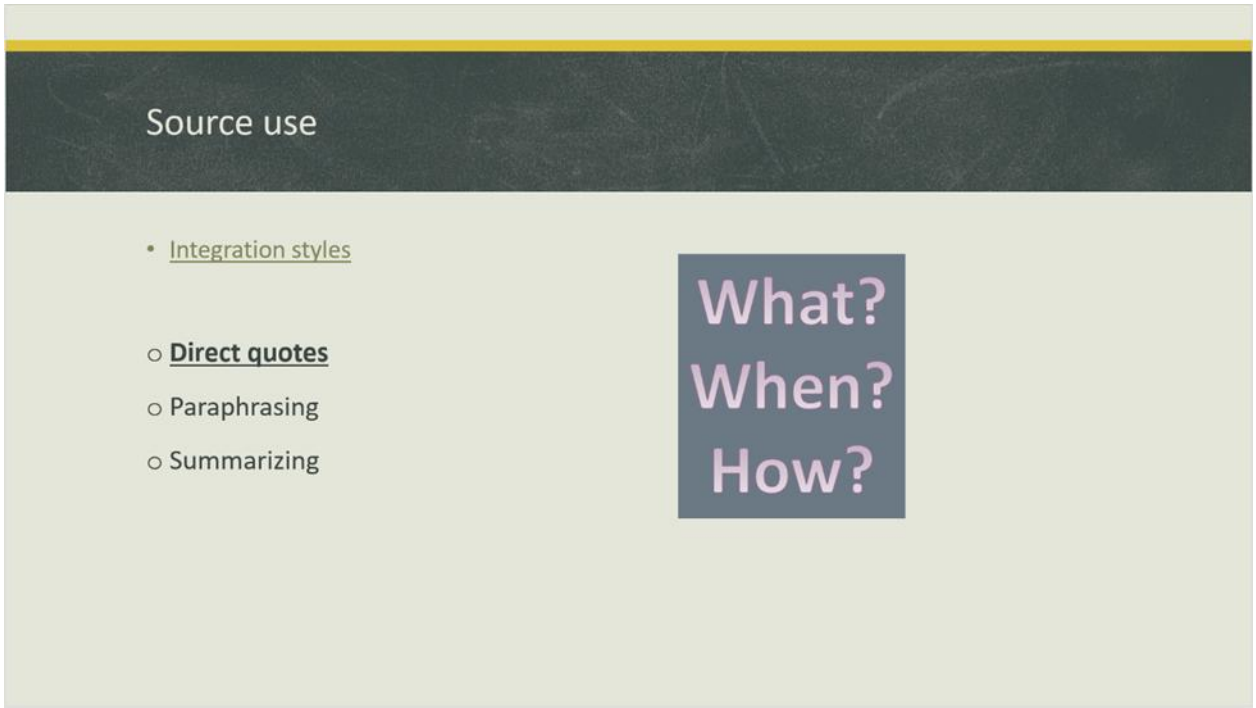
  - TOEFL iBT: Summarize the points raised in the lecture and explain how they challenge the ideas that are discussed in the reading text.
  - IELTs: The bar chart shows information about the number of times that residents and non-residents drive into the city center. Summarize this information by selecting and reporting the major features, making comparisons when relevant.
  - Lab report: Describe and graphically present your experimental data and discuss findings.

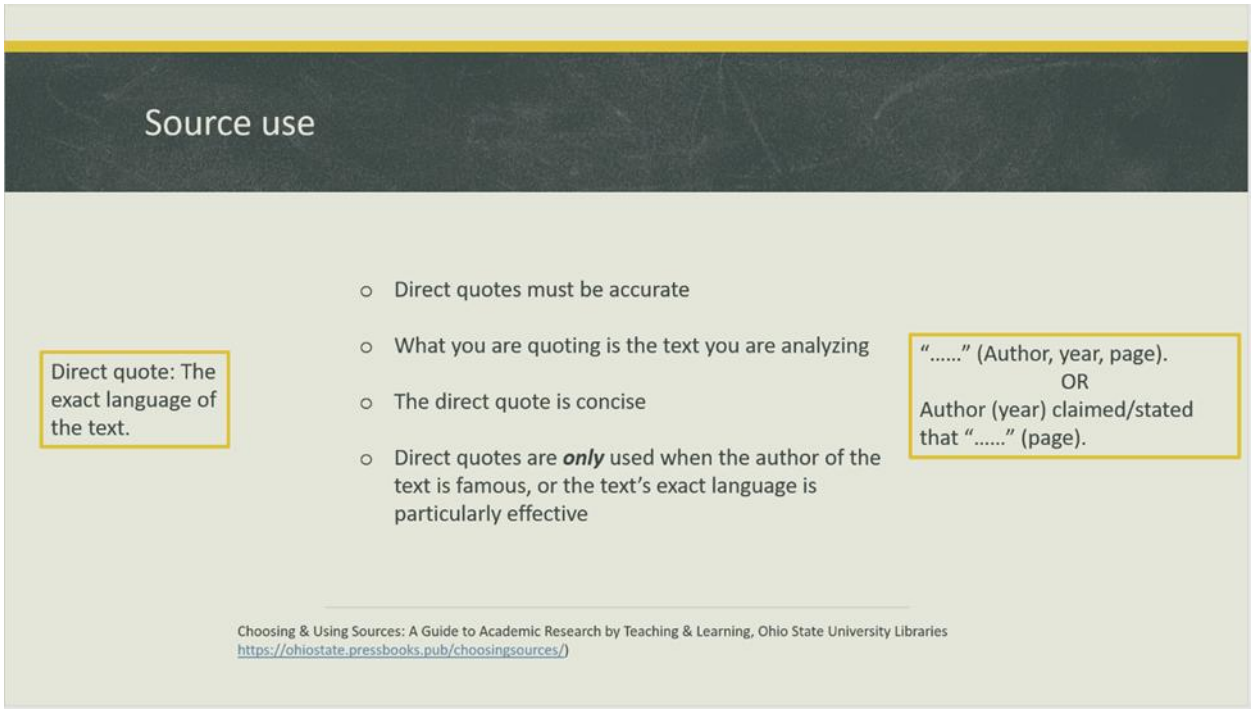


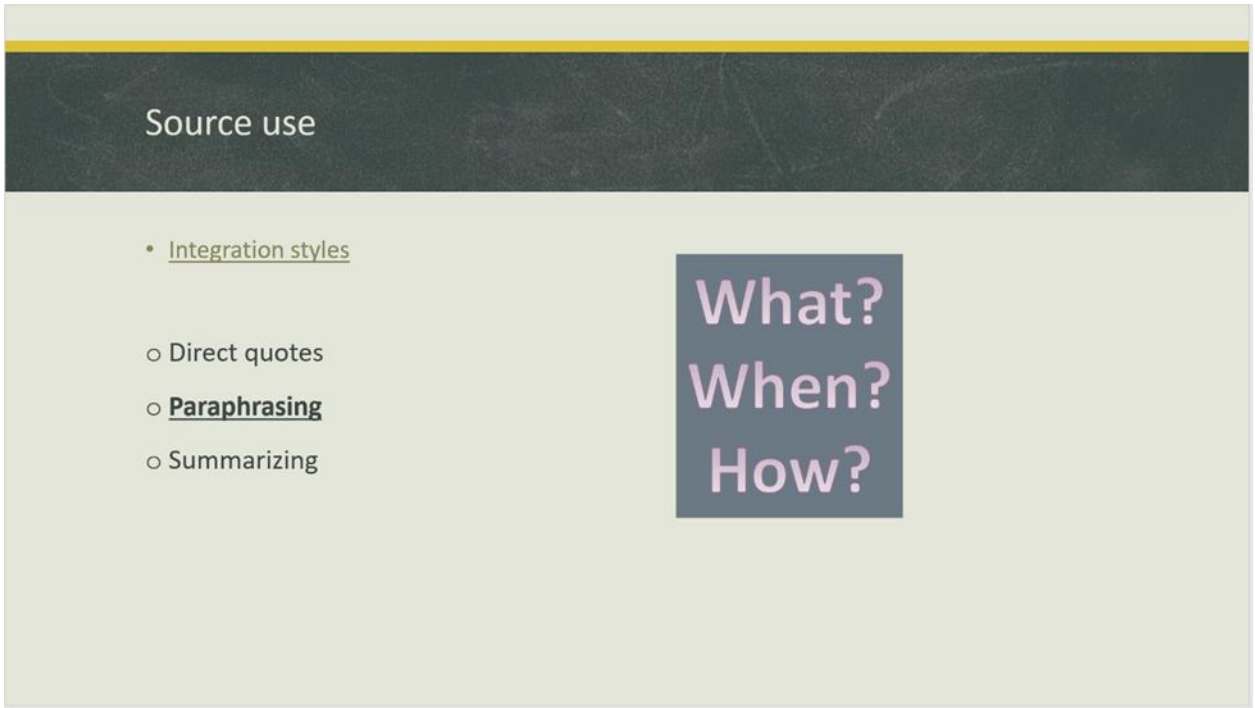

Click to add title
Instructions:
1. Make sure to use at least three points of information from the reading texts to support your opinion.
2. The essay should be about 250 words.
3. Your writing will be rated based on how effectively you incorporate reading materials successfully and refer to authors accurately
[Please type your writing below]
(Word count: )

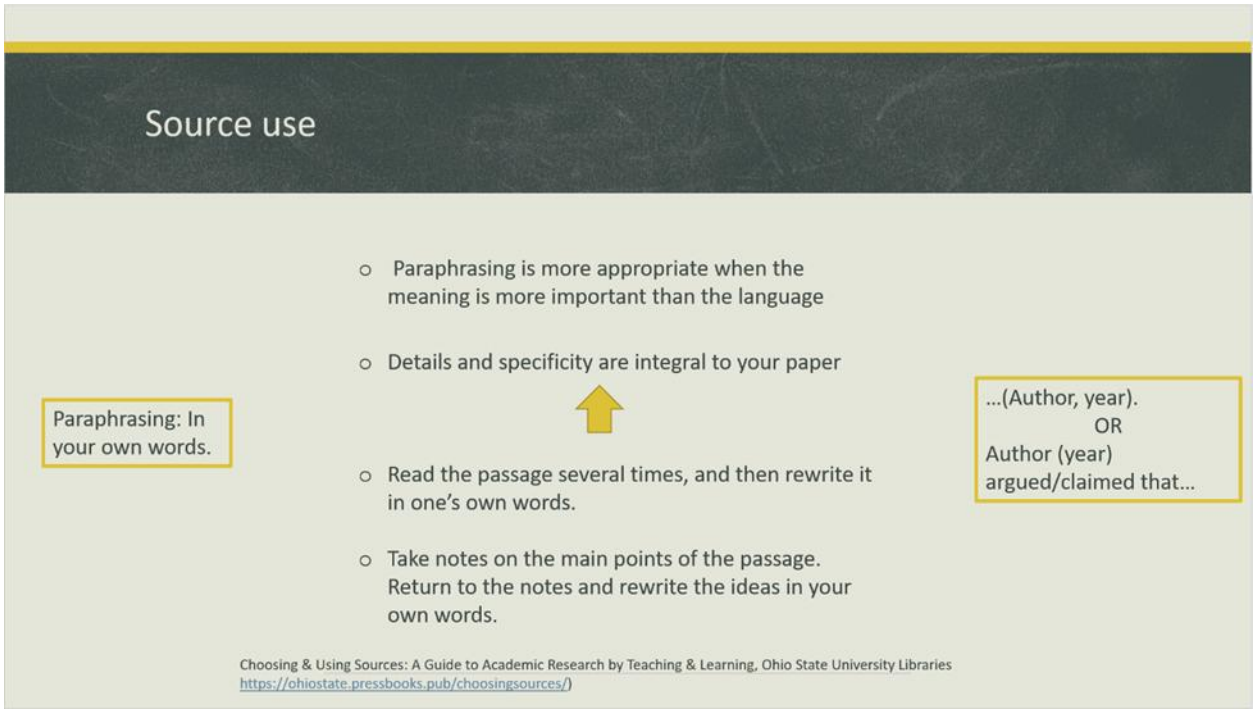

Content
Review consent form (5 mins)
Review integrated writing tasks (5 mins)
Review source use (10 mins)
Introduce the assessment checklist (5 mins)
Practice using the assessment checklist (20 mins)

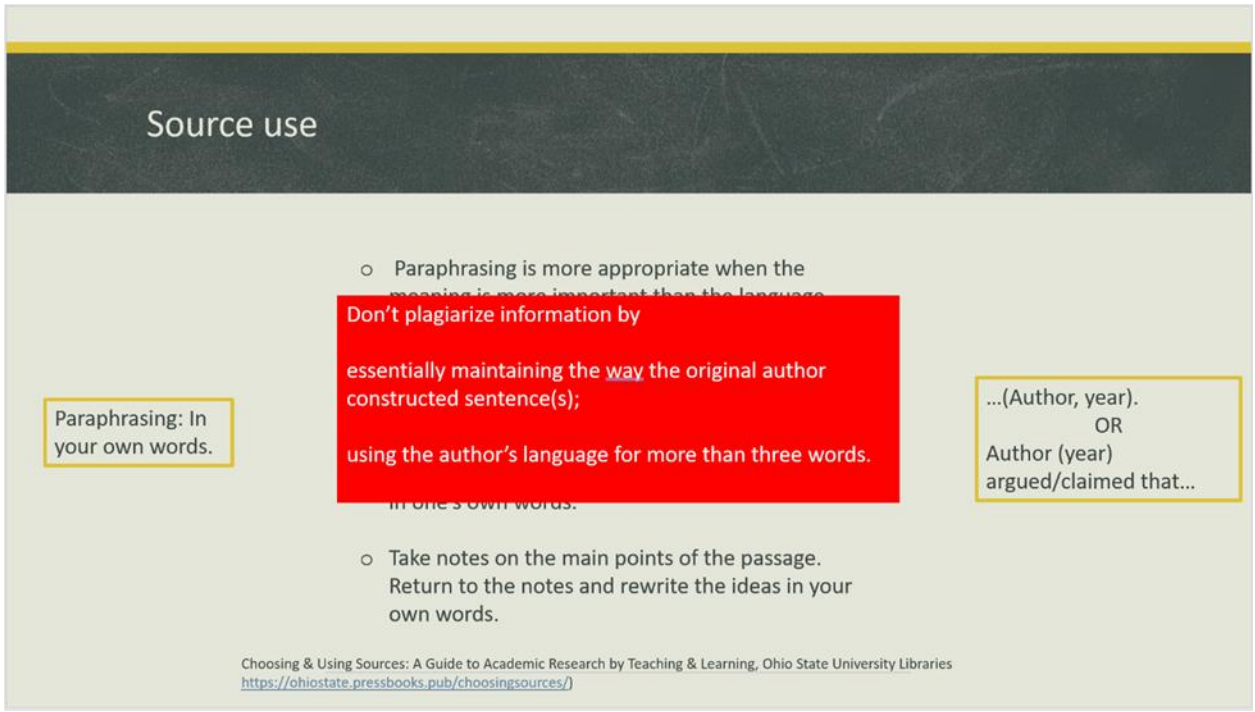

Source use
• Integration styles
○ Direct quotes
○ Paraphrasing
○ Summarizing

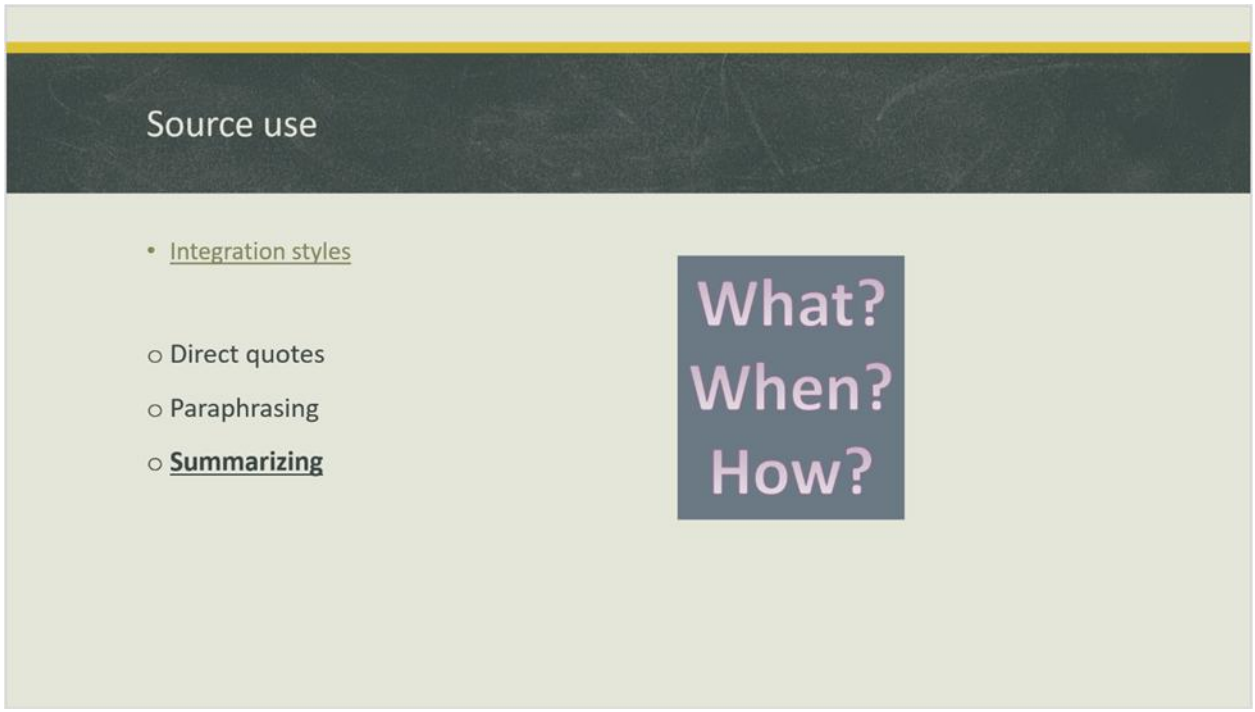

Source use
• Integration styles
○ Direct quotes
○ Paraphrasing
○ Summarizing
What?
When?
How?

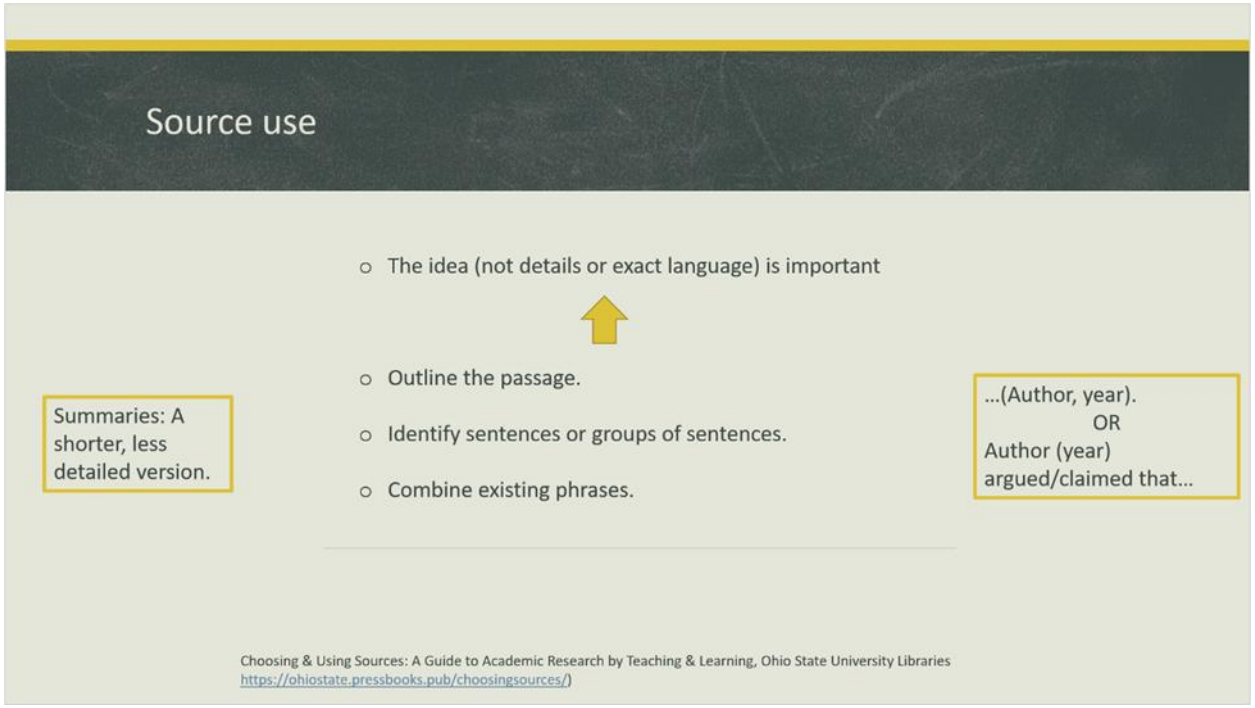

Source use
Direct quote: The exact language of the text.
Direct quotes must be accurate
What you are quoting is the text you are analyzing
The direct quote is concise
Direct quotes are only used when the author of the text is famous, or the text's exact language is particularly effective
"....." (Author, year, page).
OR
Author (year) claimed/stated that "....." (page).
Choosing & Using Sources: A Guide to Academic Research by Teaching & Learning, Ohio State University Libraries
https://ohiostate.pressbooks.pub/choosingsources/)

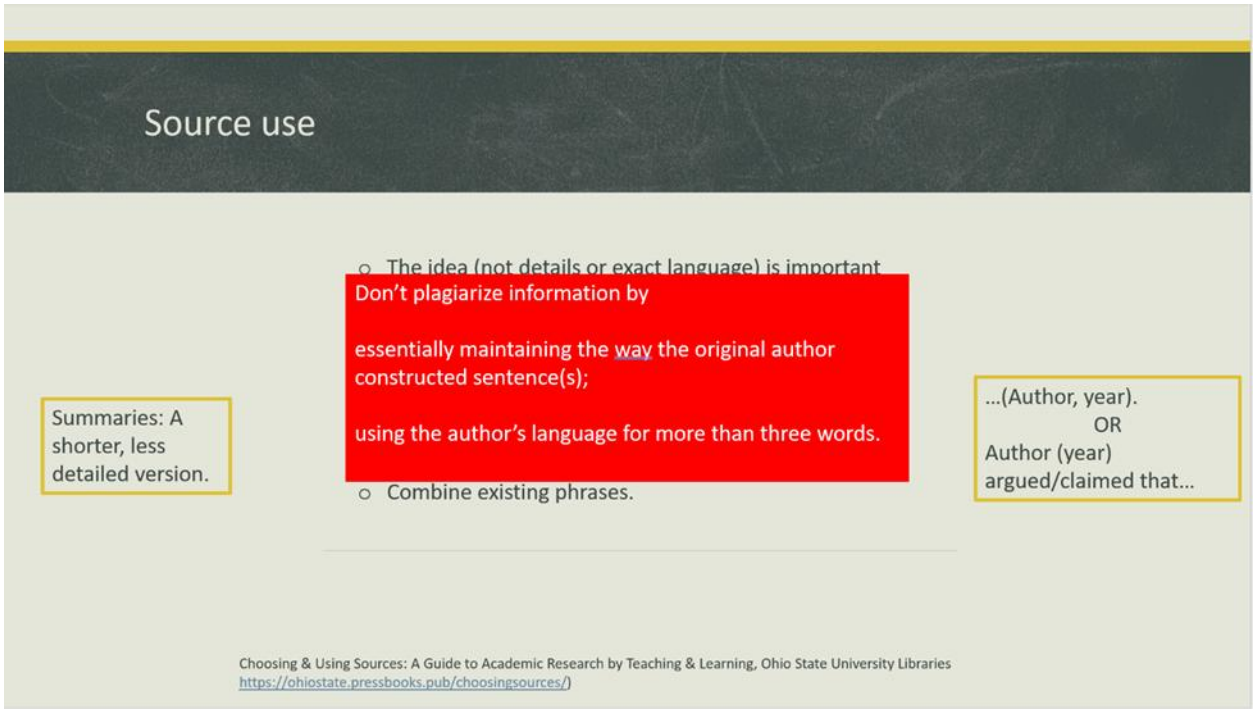

Source use
Integration styles
Direct quotes
Paraphrasing
Summarizing
What?
When?
How?

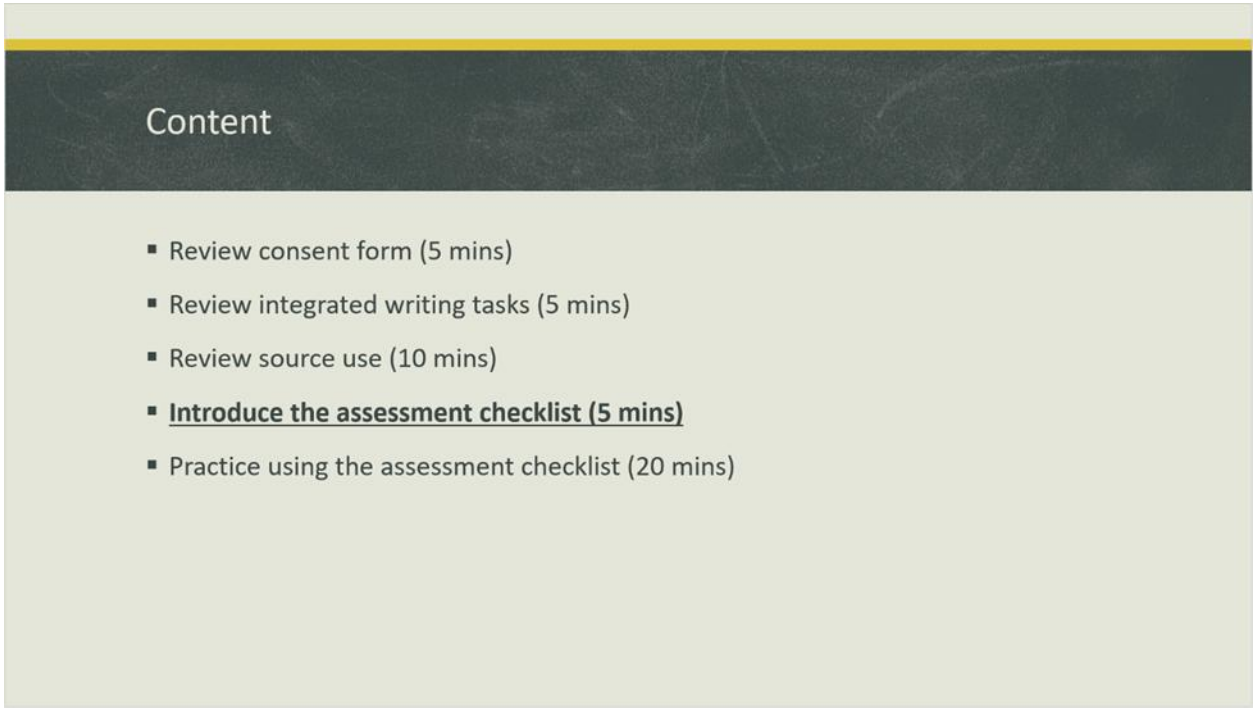

Source use
Paraphrasing is more appropriate when the meaning is more important than the language
Details and specificity are integral to your paper
Read the passage several times, and then rewrite it in one's own words.
Take notes on the main points of the passage. Return to the notes and rewrite the ideas in your own words.
Paraphrasing: In your own words.
...(Author, year).
OR
Author (year) argued/claimed that...
Choosing & Using Sources: A Guide to Academic Research by Teaching & Learning, Ohio State University Libraries
https://ohiostate.pressbooks.pub/choosingsources/)


Source use
Paraphrasing is more appropriate when the
Don't plagiarize information by
essentially maintaining the way the original author constructed sentence(s);
using the author's language for more than three words.
Take notes on the main points of the passage. Return to the notes and rewrite the ideas in your own words.
Paraphrasing: In your own words.
...(Author, year).
OR
Author (year) argued/claimed that...
Choosing & Using Sources: A Guide to Academic Research by Teaching & Learning, Ohio State University Libraries
https://ohiostate.pressbooks.pub/choosingsources/)


Source use
Integration styles
Direct quotes
Paraphrasing
Summarizing
What?
When?
How?


Source use
The idea (not details or exact language) is important
Outline the passage.
Identify sentences or groups of sentences.
Combine existing phrases.
Summaries: A shorter, less detailed version.
...(Author, year).
OR
Author (year) argued/claimed that...
Choosing & Using Sources: A Guide to Academic Research by Teaching & Learning, Ohio State University Libraries
https://ohiostate.pressbooks.pub/choosingsources/)


Source use
The idea (not details or exact language) is important
Don't plagiarize information by
essentially maintaining the way the original author constructed sentence(s);
using the author's language for more than three words.
Combine existing phrases.
Summaries: A shorter, less detailed version.
...(Author, year).
OR
Author (year) argued/claimed that...
Choosing & Using Sources: A Guide to Academic Research by Teaching & Learning, Ohio State University Libraries
https://ohiostate.pressbooks.pub/choosingsources/)


Content
Review consent form (5 mins)
Review integrated writing tasks (5 mins)
Review source use (10 mins)
Introduce the assessment checklist (5 mins)
Practice using the assessment checklist (20 mins)

## Assessment checklist

**Assessment Checklist**

Below is a brief checklist to help you with self-assessment. The checklist has three sections, but you do not need to complete each section if you did not use a particular integration style. When conducting self-assessment, please check (√) the box if you think your essay meets that criterion and place a (X) if you think your essay does not meet that criterion. You can provide some comments for yourself if needed.

If direct quotation(s) were used in your essay, please check if you met the following criterion:

□ When I directly quoted materials from the readings, I properly referred to and cited the author(s).

□ I used direct quotation(s) only when the source's exact language is particularly effective.

Comments (optional):

If paraphrasing was used in your essay, please check if you met the following criteria:

□ I properly referred to and cited the author(s).

□ I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).

□ I did not plagiarize information by using the author's language for more than three words.

Comments (optional):

If summarizing was used in your essay, please check if you met the following criteria:

□ I properly referred to and cited the author(s).

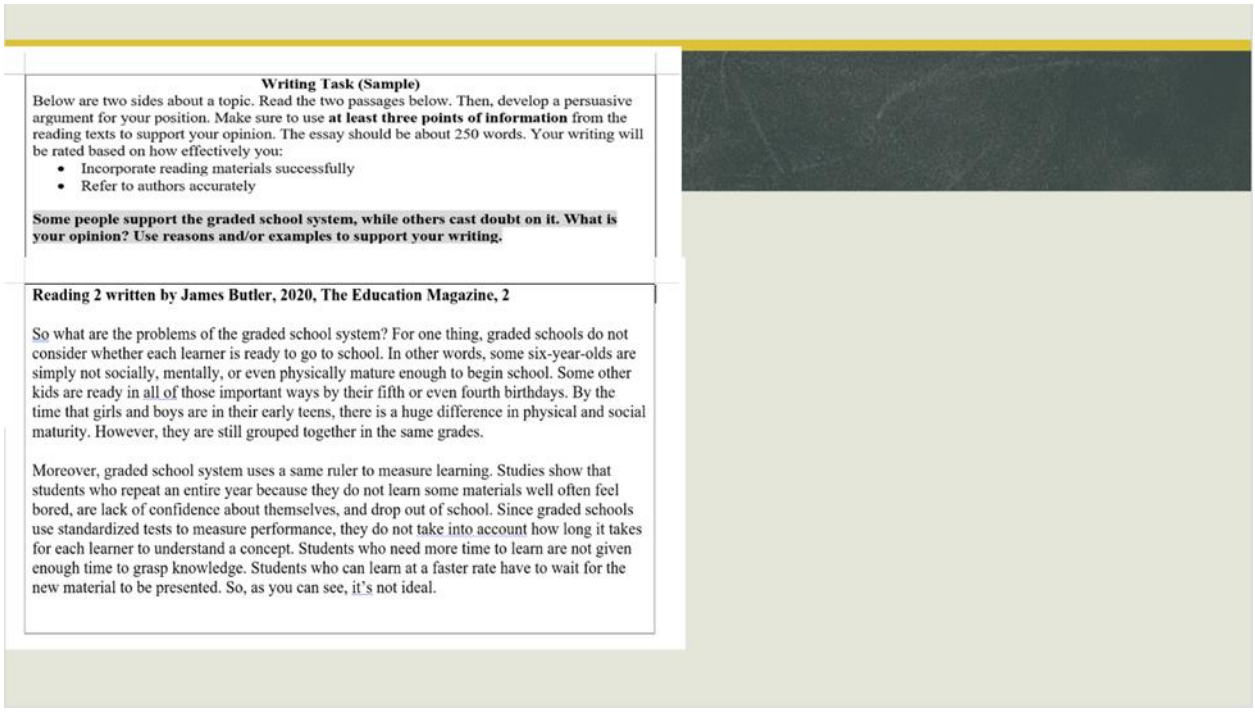
□ I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).

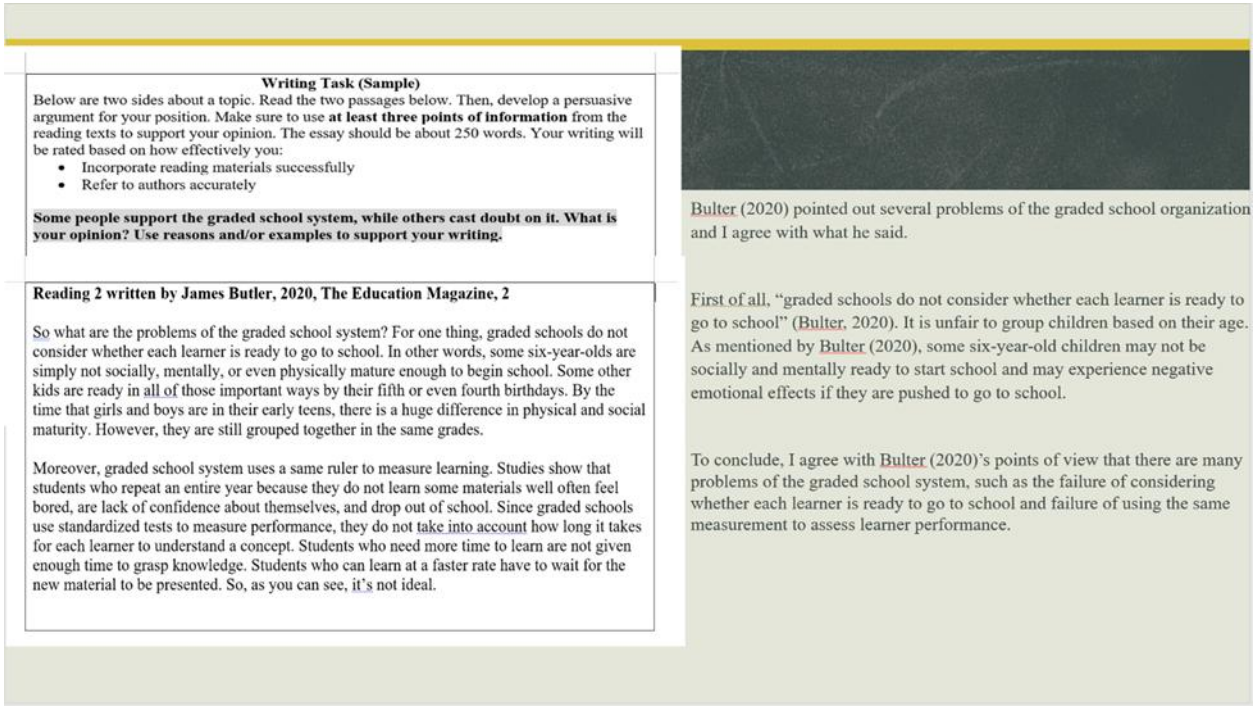
□ I did not plagiarize information by using the author's language for more than three words.

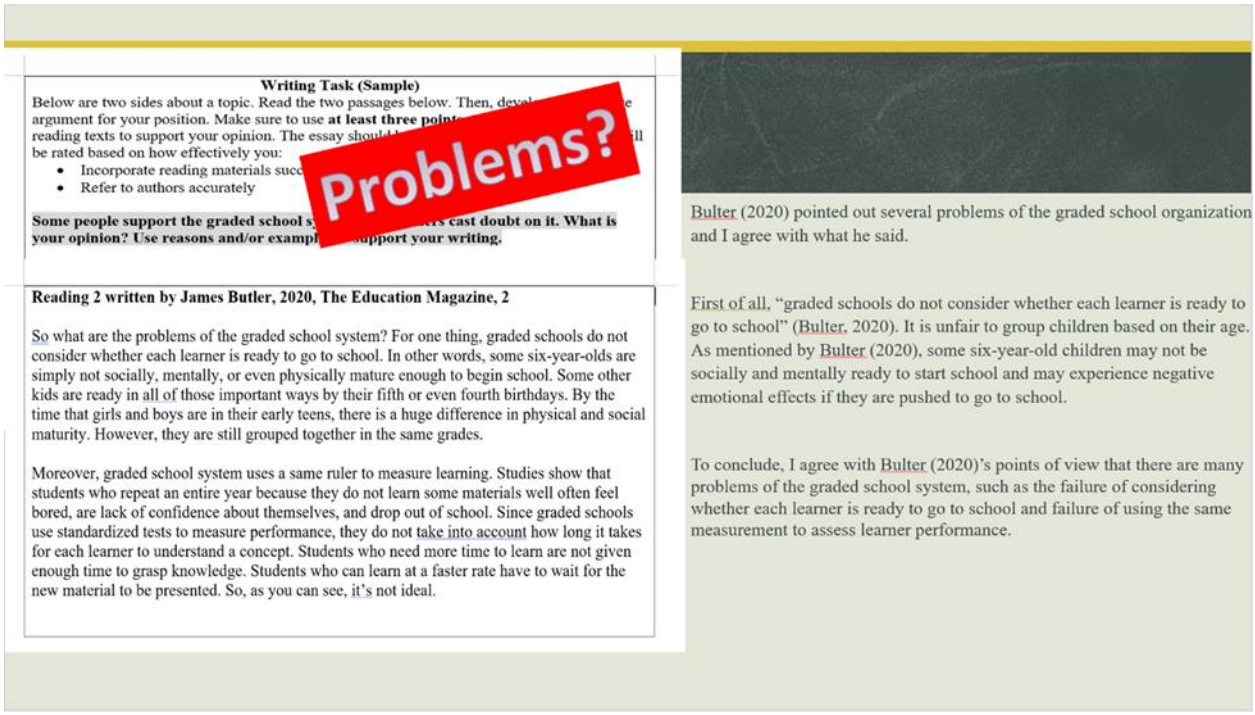
Comments (optional):

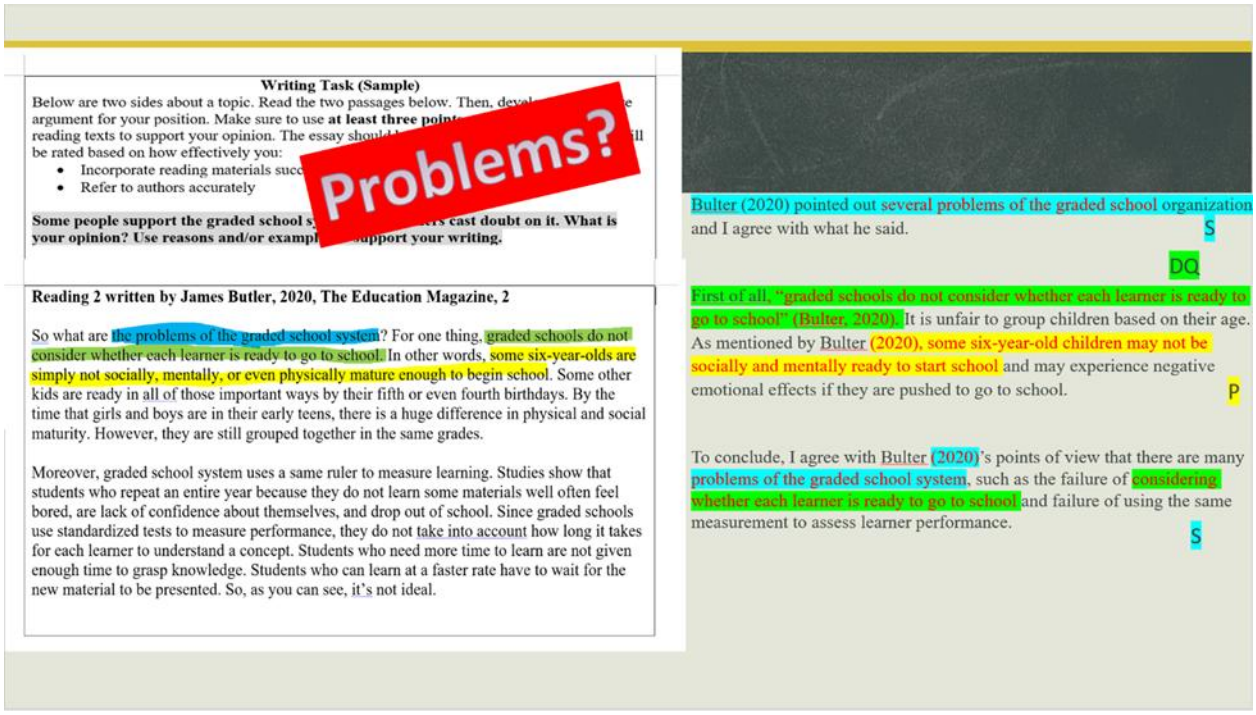
## Content

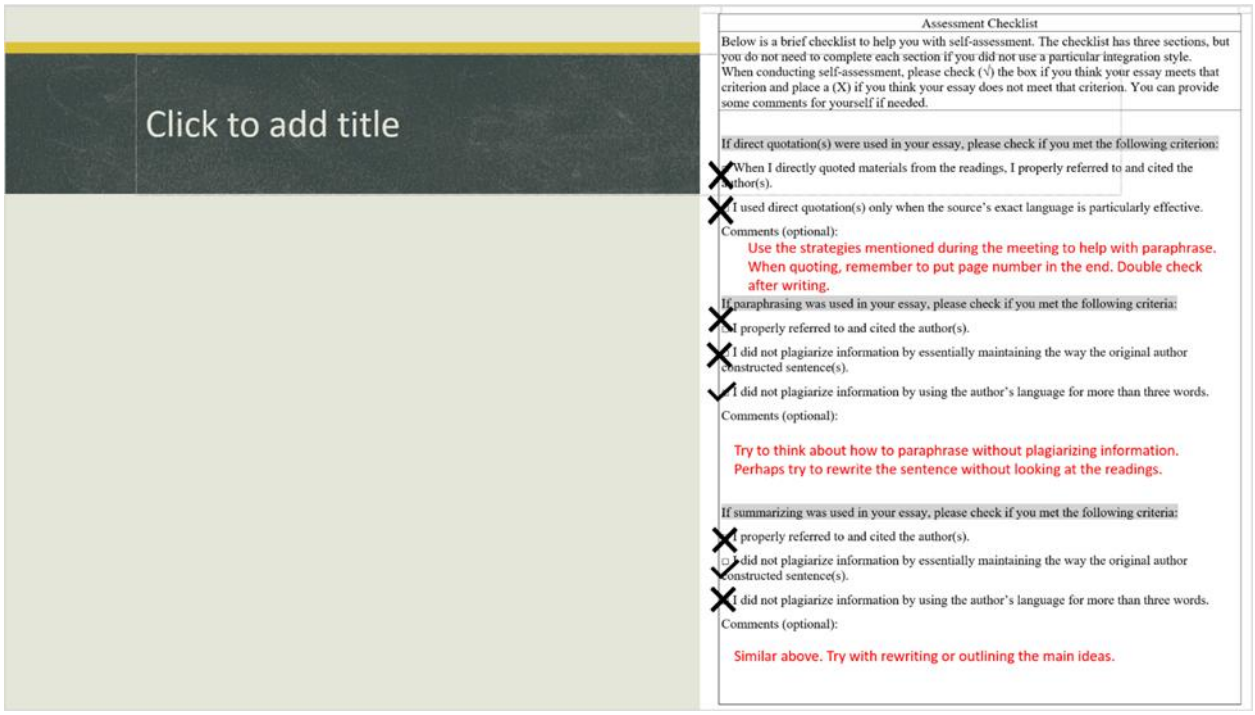
- Review consent form (5 mins)
- Introduce integrated reading-to-write tasks (5 mins)
- Introduce source use (10 mins)
- Introduce the assessment checklist (5 mins)
- **Practice using the assessment checklist (20 mins)**

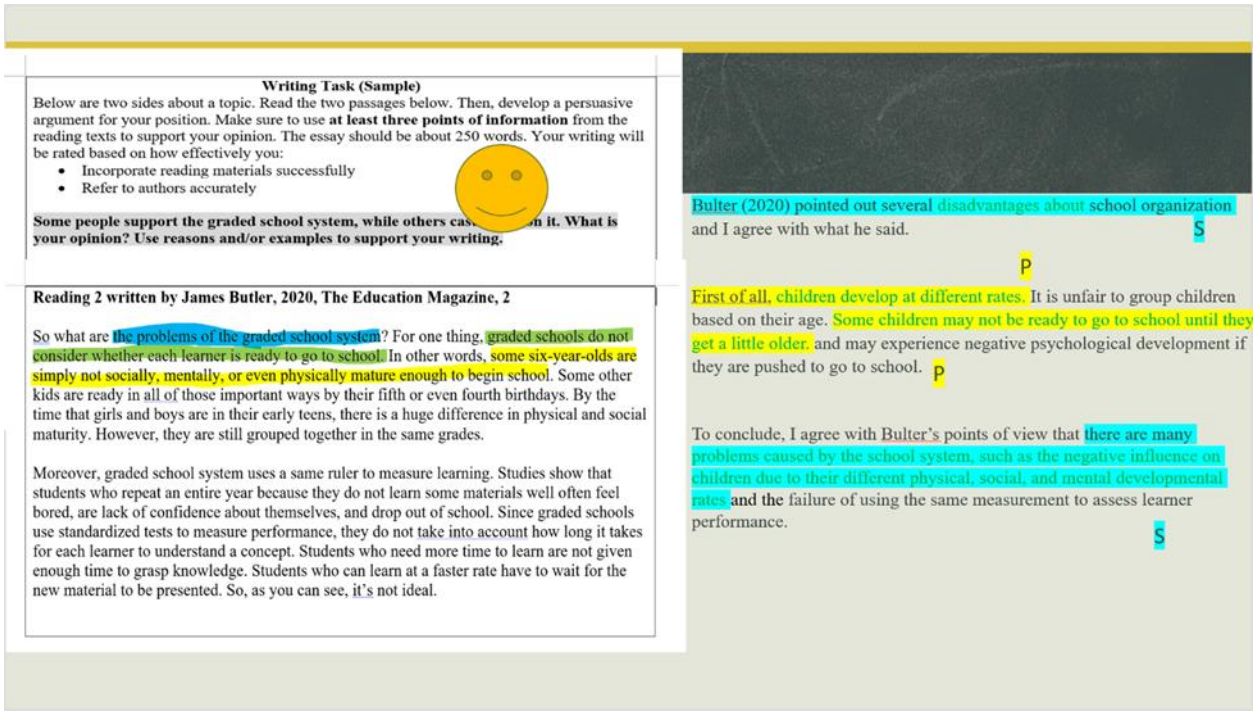

Writing Task (Sample)
Below are two sides about a topic. Read the two passages below. Then, develop a persuasive argument for your position. Make sure to use at least three points of information from the reading texts to support your opinion. The essay should be about 250 words. Your writing will be rated based on how effectively you:
Incorporate reading materials successfully
Refer to authors accurately
Some people support the graded school system, while others cast doubt on it. What is your opinion? Use reasons and/or examples to support your writing.
Reading 2 written by James Butler, 2020, The Education Magazine, 2
So what are the problems of the graded school system? For one thing, graded schools do not consider whether each learner is ready to go to school. In other words, some six-year-olds are simply not socially, mentally, or even physically mature enough to begin school. Some other kids are ready in all of those important ways by their fifth or even fourth birthdays. By the time that girls and boys are in their early teens, there is a huge difference in physical and social maturity. However, they are still grouped together in the same grades.
Moreover, graded school system uses a same ruler to measure learning. Studies show that students who repeat an entire year because they do not learn some materials well often feel bored, are lack of confidence about themselves, and drop out of school. Since graded schools use standardized tests to measure performance, they do not take into account how long it takes for each learner to understand a concept. Students who need more time to learn are not given enough time to grasp knowledge. Students who can learn at a faster rate have to wait for the new material to be presented. So, as you can see, it's not ideal.

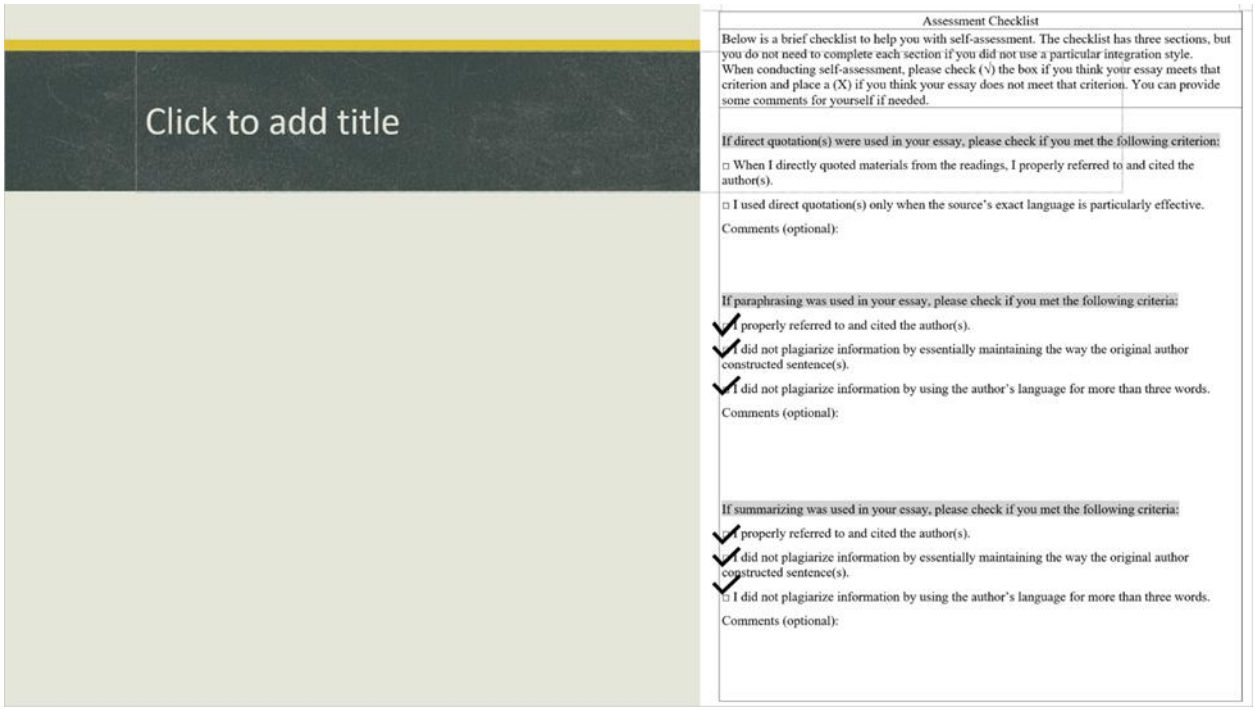

Writing Task (Sample)
Below are two sides about a topic. Read the two passages below. Then, develop a persuasive argument for your position. Make sure to use at least three points of information from the reading texts to support your opinion. The essay should be about 250 words. Your writing will be rated based on how effectively you:
Incorporate reading materials successfully
Refer to authors accurately
Some people support the graded school system, while others cast doubt on it. What is your opinion? Use reasons and/or examples to support your writing.
Reading 2 written by James Butler, 2020, The Education Magazine, 2
So what are the problems of the graded school system? For one thing, graded schools do not consider whether each learner is ready to go to school. In other words, some six-year-olds are simply not socially, mentally, or even physically mature enough to begin school. Some other kids are ready in all of those important ways by their fifth or even fourth birthdays. By the time that girls and boys are in their early teens, there is a huge difference in physical and social maturity. However, they are still grouped together in the same grades.
Moreover, graded school system uses a same ruler to measure learning. Studies show that students who repeat an entire year because they do not learn some materials well often feel bored, are lack of confidence about themselves, and drop out of school. Since graded schools use standardized tests to measure performance, they do not take into account how long it takes for each learner to understand a concept. Students who need more time to learn are not given enough time to grasp knowledge. Students who can learn at a faster rate have to wait for the new material to be presented. So, as you can see, it's not ideal.
Bulter (2020) pointed out several problems of the graded school organization and I agree with what he said.
First of all, "graded schools do not consider whether each learner is ready to go to school" (Bulter, 2020). It is unfair to group children based on their age. As mentioned by Bulter (2020), some six-year-old children may not be socially and mentally ready to start school and may experience negative emotional effects if they are pushed to go to school.
To conclude, I agree with Bulter (2020)'s points of view that there are many problems of the graded school system, such as the failure of considering whether each learner is ready to go to school and failure of using the same measurement to assess learner performance.

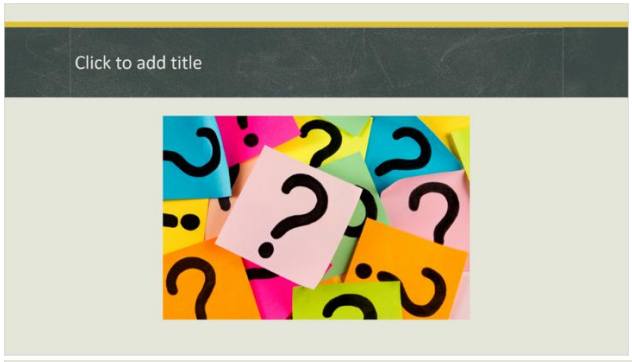

Writing Task (Sample)
Problems?
Reading 2 written by James Butler, 2020, The Education Magazine, 2
So what are the problems of the graded school system? For one thing, graded schools do not consider whether each learner is ready to go to school. In other words, some six-year-olds are simply not socially, mentally, or even physically mature enough to begin school. Some other kids are ready in all of those important ways by their fifth or even fourth birthdays. By the time that girls and boys are in their early teens, there is a huge difference in physical and social maturity. However, they are still grouped together in the same grades.
Moreover, graded school system uses a same ruler to measure learning. Studies show that students who repeat an entire year because they do not learn some materials well often feel bored, are lack of confidence about themselves, and drop out of school. Since graded schools use standardized tests to measure performance, they do not take into account how long it takes for each learner to understand a concept. Students who need more time to learn are not given enough time to grasp knowledge. Students who can learn at a faster rate have to wait for the new material to be presented. So, as you can see, it's not ideal.
Bulter (2020) pointed out several problems of the graded school organization and I agree with what he said.
First of all, "graded schools do not consider whether each learner is ready to go to school" (Bulter, 2020). It is unfair to group children based on their age. As mentioned by Bulter (2020), some six-year-old children may not be socially and mentally ready to start school and may experience negative emotional effects if they are pushed to go to school.
To conclude, I agree with Bulter (2020)'s points of view that there are many problems of the graded school system, such as the failure of considering whether each learner is ready to go to school and failure of using the same measurement to assess learner performance.

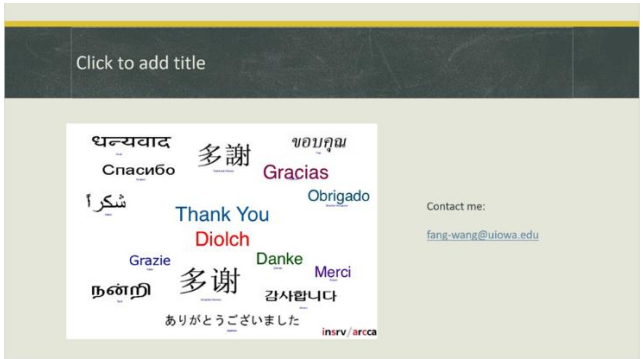

Writing Task (Sample)
Problems?
Reading 2 written by James Butler, 2020, The Education Magazine, 2
So what are the problems of the graded school system? For one thing, graded schools do not consider whether each learner is ready to go to school. In other words, some six-year-olds are simply not socially, mentally, or even physically mature enough to begin school. Some other kids are ready in all of those important ways by their fifth or even fourth birthdays. By the time that girls and boys are in their early teens, there is a huge difference in physical and social maturity. However, they are still grouped together in the same grades.
Moreover, graded school system uses a same ruler to measure learning. Studies show that students who repeat an entire year because they do not learn some materials well often feel bored, are lack of confidence about themselves, and drop out of school. Since graded schools use standardized tests to measure performance, they do not take into account how long it takes for each learner to understand a concept. Students who need more time to learn are not given enough time to grasp knowledge. Students who can learn at a faster rate have to wait for the new material to be presented. So, as you can see, it's not ideal.
Bulter (2020) pointed out several problems of the graded school organization and I agree with what he said.
S
DQ
First of all, "graded schools do not consider whether each learner is ready to go to school" (Bulter, 2020). It is unfair to group children based on their age. As mentioned by Bulter (2020), some six-year-old children may not be socially and mentally ready to start school and may experience negative emotional effects if they are pushed to go to school.
P
To conclude, I agree with Bulter (2020)'s points of view that there are many problems of the graded school system, such as the failure of considering whether each learner is ready to go to school and failure of using the same measurement to assess learner performance.
S


Click to add title
Assessment Checklist
Below is a brief checklist to help you with self-assessment. The checklist has three sections, but you do not need to complete each section if you did not use a particular integration style. When conducting self-assessment, please check (√) the box if you think your essay meets that criterion and place a (X) if you think your essay does not meet that criterion. You can provide some comments for yourself if needed.
If direct quotation(s) were used in your essay, please check if you met the following criterion:
When I directly quoted materials from the readings, I properly referred to and cited the author(s).
I used direct quotation(s) only when the source's exact language is particularly effective.
Comments (optional):
Use the strategies mentioned during the meeting to help with paraphrase. When quoting, remember to put page number in the end. Double check after writing.
If paraphrasing was used in your essay, please check if you met the following criteria:
I properly referred to and cited the author(s).
I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).
I did not plagiarize information by using the author's language for more than three words.
Comments (optional):
Try to think about how to paraphrase without plagiarizing information. Perhaps try to rewrite the sentence without looking at the readings.
If summarizing was used in your essay, please check if you met the following criteria:
I properly referred to and cited the author(s).
I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).
I did not plagiarize information by using the author's language for more than three words.
Comments (optional):
Similar above. Try with rewriting or outlining the main ideas.


Writing Task (Sample)
Below are two sides about a topic. Read the two passages below. Then, develop a persuasive argument for your position. Make sure to use at least three points of information from the reading texts to support your opinion. The essay should be about 250 words. Your writing will be rated based on how effectively you:
Incorporate reading materials successfully
Refer to authors accurately
Some people support the graded school system, while others cas... it. What is your opinion? Use reasons and/or examples to support your writing.
Reading 2 written by James Butler, 2020, The Education Magazine, 2
So what are the problems of the graded school system? For one thing, graded schools do not consider whether each learner is ready to go to school. In other words, some six-year-olds are simply not socially, mentally, or even physically mature enough to begin school. Some other kids are ready in all of those important ways by their fifth or even fourth birthdays. By the time that girls and boys are in their early teens, there is a huge difference in physical and social maturity. However, they are still grouped together in the same grades.
Moreover, graded school system uses a same ruler to measure learning. Studies show that students who repeat an entire year because they do not learn some materials well often feel bored, are lack of confidence about themselves, and drop out of school. Since graded schools use standardized tests to measure performance, they do not take into account how long it takes for each learner to understand a concept. Students who need more time to learn are not given enough time to grasp knowledge. Students who can learn at a faster rate have to wait for the new material to be presented. So, as you can see, it's not ideal.
Bulter (2020) pointed out several disadvantages about school organization and I agree with what he said.
S
P
First of all, children develop at different rates. It is unfair to group children based on their age. Some children may not be ready to go to school until they get a little older. and may experience negative psychological development if they are pushed to go to school.
P
To conclude, I agree with Bulter's points of view that there are many problems caused by the school system, such as the negative influence on children due to their different physical, social, and mental developmental rates and the failure of using the same measurement to assess learner performance.
S


Click to add title
Assessment Checklist
Below is a brief checklist to help you with self-assessment. The checklist has three sections, but you do not need to complete each section if you did not use a particular integration style. When conducting self-assessment, please check (√) the box if you think your essay meets that criterion and place a (X) if you think your essay does not meet that criterion. You can provide some comments for yourself if needed.
If direct quotation(s) were used in your essay, please check if you met the following criterion:
□ When I directly quoted materials from the readings, I properly referred to and cited the author(s).
□ I used direct quotation(s) only when the source's exact language is particularly effective.
Comments (optional):
If paraphrasing was used in your essay, please check if you met the following criteria:
√ I properly referred to and cited the author(s).
√ I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).
√ I did not plagiarize information by using the author's language for more than three words.
Comments (optional):
If summarizing was used in your essay, please check if you met the following criteria:
√ I properly referred to and cited the author(s).
√ I did not plagiarize information by essentially maintaining the way the original author constructed sentence(s).
□ I did not plagiarize information by using the author's language for more than three words.
Comments (optional):


Click to add title


Click to add title
धन्यवाद
多謝
ขอบคุณ
Спасибо
Gracias
شكرا
Obrigado
Thank You
Diolch
Grazie
Danke
Merci
நன்றி
多谢
감사합니다
ありがとうございました
insrv / arcca
Contact me:
fang-wang@uiowa.edu

**Appendix F.** *Purpose of the pilot study.*

| | | **Study participants** |
|---|---|---|
| Participants | Appropriateness of the subjects | Can participants benefit from the study given their English language proficiency?<br>AND<br>Can participants benefit from the study given their experience with integrated writing essays? |
| | Number of subjects | What is the attrition rate?<br>AND<br>Is the number of subjects enough for the study after having some subjects withdrawn? |
| | | **Instruments** |
| Introductory meeting | Validity of the presentation | Are the explanations of direct quotations, paraphrasing, and summarizing aligned well with the descriptors on the assessment checklist?<br>AND<br>Are the explanations of references aligned well with the descriptors on the assessment checklist? |
| | Clarity of the presentation | Can participants understand the study procedure and requirements?<br>AND<br>Can participants understand when and how to use direct quotations, paraphrasing, and summarizing?<br>AND<br>Can participants understand how to cite authors?<br>AND<br>Can participants understand the writing requirements?<br>AND<br>Can participants understand how to use the assessment checklist? |
| | Mode of the presentation | Can participants interact with me often and when needed? |
| | Length of the presentation | Is the length of the presentation proper to cover key information while not making participants feel fatigued?<br>AND<br>Is the length of the presentation completed within the planned time? |
| | Speed of the presentation | Do I speak at a moderate speed? |
| | Language of the presentation | Should I speak in Chinese to explain/clarify key information? |
| Writing tasks | Organization of the writing tasks | Is the organization clear and easy to follow? |
| | Understanding of the instructions | Can participants understand the writing requirements?<br>AND<br>Are all requirements conveyed to participants? |
| | Difficulty level of the texts | Is the difficulty level of the texts moderate? |
| | Language of the tasks | Do the writing tasks contain terms that participants may not be able to understand? |
| | Length of the texts | Is the length of the texts moderate? |
| | Topic familiarity of the texts | Is topic effect a moderate factor in affecting participants' writing? |
| Essay paper | Understanding of the essay paper | Do participants know how to use essay paper to write essays?<br>AND<br>Do participants know how to record the word account in the end? |

| | | |
|---|---|---|
| Assessment checklist | Understanding of the instructions | Are the assessment instructions clear and easy to understand?<br>AND<br>Are all instructions provided for participants? |
| | Organization of the checklist | Is the checklist organized clearly? |
| | Validity of the checklist | Do the descriptors accurately and explicitly reflect the constructs being measured?<br>AND<br>Are the descriptors transparent so that participants can base their judgements on the descriptors? |
| | Language of the checklist | Are the descriptors transparent and concise?<br>AND<br>Does the checklist contain terms that participants may not be able to understand? |
| | Practicality of the checklist | Is it easy for participants to check/uncheck the box? |
| Post-task interview | Clarity of the questions | Are the questions easy to understand?<br>AND<br>Are the questions repetitive? |
| | Validity of the questions | Are the questions able to answer the research questions?<br>AND<br>Do the questions probe/elicit as much information as possible from participants? |
| | Sequence of the questions | Are the questions put in proper order? |
| | Appropriateness of the questions | Do the questions pose any threat or discomfort to participants? |
| | | **Study design** |
| Study design | Study pace | Is the study pace moderate? |
| | Duration of the study | Is the duration of the study moderate? |
| | Reminder from the researcher | Are the reminders to participants sent timely? |

**Appendix G**. *My relfections and participants' feedback about the pilot study.*

| | Problems | Thoughts | Solutions |
|---|---|---|---|
| Lack of motivation/High attrition rates | It took longer for participants to send back the essays | Since participants were in their home country, they might have lacked the motivation to enhance their English academic writing skills | Recruit more participants |
| | Some participants withdrew midway through the study | | Send regular reminders to participants who take longer to complete essays |
| | | | Give them more time to work on the tasks |
| Lack of learning expected from the study | Participants demonstrated strong writing skills, making revisions unnecessary | It's possible that learning extended beyond improvements in writing performance, such as gaining a better understanding of plagiarism or acquiring knowledge about the content | Target low-intermediate level instead of high-intermediate level L2 writers |
| | | Maybe they were too skilled | Recruit participants who are new to academic writing |
| Difficulties in tracking changes | Participants might have revised their essays, but their changes were difficult to identify | Since many of the revisions were minor, it was somewhat challenging to identify the changes participants had made | Ask participants to use the tracking function during revisions, as this could potentially enhance the likelihood of making meaningful revisions. |
| | | Participants may not revise the essay at all | |
| Wrong genre | One participant wrote two summaries of the two reading passages instead of one argumentation | The participant may not know what an argumentation is | Check their understanding at the introductory meeting |
| | | The participant may forget the writing requirement after reading | Include the writing requirement and prompt at the top of the essay paper |
| No citation style used in their writing | Some participants had not been exposed to any citation style | Since the participants had only recently begun their ESL writing class and had not yet taken any disciplinary courses, they were unfamiliar with any specific citation styles. | Ask about their program of study and check the citation style they will use in the future |
| | | | Can introduce the APA style to them if they are not certain about their major |