

The Comparative Efficacy of Instruction of Source Text Use in Reading-to-write versus Listening-to-write Integrated Writing Tasks

¹Maryam Azarian
²Mahboobeh Saadat*

Research Paper

IJEAP-2602-2188

Received: 2026-02-03

Accepted: 2026-07-07

Published: 2026-09-30

Abstract: A growing body of research addresses a variety of issues regarding Integrated Writing (IW), used to assess learners' academic writing ability in different contexts. However, the effect of source use strategy instruction on learners' IW ability is still under-researched. To address this niche, the present study, with a pre-experimental one-group pretest-posttest-delayed posttest design, compared the effect of an instructional intervention on the improvement of English as a Foreign Language (EFL) learners' source use ability in reading-to-write (RW) and listening-to-write (LW) tasks. To this end, the required data were mainly accumulated from the written tasks of 50 Iranian EFL sophomores studying in two intact classes of Essay Writing, comprising one experimental group, during eight sessions. Both classes had the same instructor and received identical intervention concerning strategies of appropriate source employment in RW and LW tasks assigned to them alternatively with different order across the two classes. The analyses of quantitative and qualitative data may suggest that the instructional intervention could significantly improve participants' source-based writing, and the contribution of the intervention to the improvement of their RW and LW tasks was similar. Moreover, from among textual borrowing patterns, non-referenced reformulated and modified source materials, and direct quotation were employed more frequently than the others in both tasks. Findings also illuminated the challenging nature of source integration and borrowing, specifically in LW tasks. Finally, pedagogical implications were discussed with a primary emphasis on the need to include the necessary instructional intervention on synthesis writing and source use strategies in writing classes.

Keywords: EFL Learners, Instructional Intervention, Integrated Writing, Listening-to-write, Reading-to-write, Source Use

Introduction

Integrated Writing (IW) tasks have been widely employed to measure Second Language (L2) learners' writing ability and have caught the attention of a large number of researchers (e.g., Cai & Yan, 2024; Cumming et al., 2016; Gebril & Plakans, 2013; Golparvar et al., 2025; Grabe & Zhang, 2013a, 2013b; Plakans, 2008; Xu et al., 2025). In doing IW tasks, learners incorporate notions from other sources, specifically reading and listening sources, into their writing products (e.g., Cooney et al., 2018; Grabe & Zhang, 2013a, 2013b; Plakans & Gebril, 2013), leading to intermingling of their receptive and productive skills. Indeed, these tasks differ from independent writing tasks, which rely solely on learners' prior knowledge (Guo et al., 2013), in the integration of source materials (Plakans & Gebril, 2013). IW tasks, through the employment of oral, written or multimodal sources as stimuli, get close to academic writing tasks (Cumming et al., 2005) which usually react to and are dependent on other pieces of writing (Karimnia & Safnyyat, 2021). Hence, as Leki and Carson (1997) argued, writing in English for Academic Purposes (EAP) should shift from personal opinion-based tasks to those integrating

¹ PhD Candidate of TEFL, Email: mazarian2011@gmail.com; Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran.

² Associate Professor of TEFL (Corresponding Author), Email: msaadat@rose.shirazu.ac.ir; Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran.

external sources because while source texts offer scaffolding, they also introduce complexity. Hyland (2009) also stressed the importance of appropriate use of source materials in academic writing. Thus, as “an important aspect of success in higher education” (Kyle, 2020, p. 1), source-based writing tasks are widely used in academic writing tasks, which are usually stimulated by outside real-world sources (e.g., Leki & Carson, 1997).

In fact, source-based writing supports learners by reducing the influence of background knowledge on their performance (Weigle, 2002), helps them generate ideas and shape attitudes, and provides language support (Plakans & Gebril, 2012). However, L2 learners’ success in producing appropriate source-based writing tasks depends on their comprehension of the aural and/or written source text(s) they are exposed to, since there is a strong connection between their writing skill and receptive skills (e.g., Cheong et al., 2018; Cumming, 2013; Gebril & Plakans, 2013). These learners are also required to apply adequate linguistic and conceptual transformation in the information from different sources in order to create a completely novel text (Boscolo et al., 2007), and avoid replicating the related information (Plakans, 2008). Hence, due to the complex and transformative demands of IW tasks, explicit instructional intervention seems to be essential for scaffolding learners to develop writing skills necessary to fulfil such tasks, as targeted instructional interventions have been indicated to significantly affect the betterment of academic writing skills of students (e.g., Aghaalkhani et al., 2022; Cooney et al., 2018; Machili et al., 2019; Zhang, 2013).

On the other hand, several studies confirm the interdependence of reading, listening, speaking and writing and the supporting role of practice in one in the development of the others (Peregoy & Boyle, 2009). In addition, two separate lines of inquiry support the interrelationship between reading and writing (e.g., Aktaş et al., 2023; Habibi et al., 2015; Linuwih & Winardi, 2020; Zhu et al., 2025) on the one hand and listening and writing (e.g., Brito, 2015; Kamal et al., 2021) on the other. Consequently, assessments in EAP increasingly combine writing with other skills to enhance task validity and authenticity (Cumming et al., 2005; Plakans & Gebril, 2013; Weigle & Parker, 2012), offering a more comprehensive evaluation of L2 writing (Cumming, 2013; Grabe & Zhang, 2013a).

Many empirical studies investigating IW tasks have focused on source text employment from different perspectives, including the influence of proficiency levels, content provision, task complexity, writing strategies and independent language skills on participants’ source-based writing performance (e.g., Cai & Yan, 2024; Cumming et al., 2005; Golparvar & Rashidi, 2021; Jung, 2020; Soleimani & Mahdavi-pour, 2014; Uludag et al., 2019; Xu et al., 2025), the comparison of the discourse constructed in IW tasks with the one composed in independent writing tasks (Cumming et al., 2005), the comparative effect of different source materials on IW performance of the participants (e.g., Cheong et al., 2018; Plakans & Gebril, 2013; Soleimani & Mahdavi-pour, 2014), and the cognitive demand of synthesis writing and the composing processes involved in IW tasks (Golparvar & Rashidi, 2021; Plakans & Gebril, 2012). Though some surveys have emphasized the value of instruction that integrates reading and writing (e.g., Cho & Brutt-Griffler, 2015; Cooney et al., 2018; Corden, 2007; Liu et al., 2019), little research has highlighted the positive impact of listening skills instruction on the participants’ writing improvement (Berninger, 2000). Furthermore, just a few investigations (e.g., Cooney et al., 2018; Jung, 2020; Machili et al., 2019; Zhang, 2013) have probed the effect of instructional intervention focusing on the strategies of appropriate source text integration on IW performance of the participants. Therefore, this study intends to contribute to the line of enquiry by quantitatively and qualitatively examining the comparative effect of instruction of source text employment strategies on IW performance of Iranian EFL learners across reading-to-write (RW) and listening-to-write (LW) tasks.

This study draws on Skill Acquisition Theory (SAT), which views language learning as a progression from declarative knowledge (knowing about a skill) to procedural knowledge (knowing how to use it), culminating in automatization through meaningful practice (DeKeyser, 2020; Lyster & Sato, 2013; VanPatten & Benati, 2010). Automatization involves faster, more accurate, and less effortful performance (DeKeyser & Criado-Sánchez, 2012; Pawlak, 2011), requiring both practice and instructional scaffolding. In this study, explicit instruction on source text use (declarative knowledge), practice, instructor feedback and the introduction of IW scoring rubric (scaffolding) are

expected to help learners proceduralize and eventually automatize IW skills (DeKeyser, 2020; Lim & Godfroid, 2014; Lyster & Sato, 2013).

Literature Review

Many studies have examined IW tasks, focusing on how source texts enhance writing quality with an emphasis on the interdependence of comprehension and production. Some of these studies have focused on the employment of reading source materials in IW tasks, and relevant topics, whereas some others have explored issues pertinent to reading-listening-writing (RLW) or more generally multimodal IW tasks. The following sections provide a brief account of these studies, respectively.

Studies on RW

Plakans (2008) compared RW integrated and writing-only tasks using think-aloud protocols and interviews with 10 L2 students. She found that integrated tasks prompted more interactive composing processes, especially among students with greater writing experience. Plakans and Gebril (2012) used a mixed-methods approach to study Arabic female undergraduates' RW performance. Think-aloud protocols, questionnaires, and interviews revealed that source texts helped participants generate ideas and expand linguistic repertoire. A positive correlation was also found between reading comprehension and writing scores, particularly at lower levels. Weigle and Parker (2012) examined how ESL learners at different proficiency levels borrowed from source texts in RW integrated tasks. Using Shi's (2004) coding scheme, they found limited use of source materials and few significant differences across education levels, proficiency, or topics. Zhang (2013) focused on the way instruction may affect intermediate Arabic and Chinese ESL students' synthesis writing. The experimental group received targeted instruction on discourse synthesis (i.e., RW), while the control group practiced general reading and writing. Results showed greater improvement in the experimental group's synthesis writing. Neumann et al. (2019) explored how Canadian EAP students used sources in RW tasks and the challenges they faced. Analysis of essays, notes, and interviews with students and instructors revealed that students often paraphrased or summarized source texts and used them to generate ideas on unfamiliar topics. The study emphasized the need for pedagogical support to help students use sources effectively. Jung (2020) studied Korean undergraduates to examine how additional content provision affected IW and incidental vocabulary learning. Participants read an opinion essay and wrote a counter-argument, with one group receiving extra content support. Results showed that unsupported learners hesitated more and used simpler language but learned target pseudowords more effectively, suggesting deeper cognitive engagement. Golparvar and Rashidi (2021) explored task complexity in source-based writing among Iranian TEFL students. One group summarized a single text, while the other synthesized three. Findings showed that complex tasks increased cognitive demand, affecting cohesion, lexical diversity, and syntactic complexity, but reduced accuracy. Lee (2023) used log-linear cognitive diagnostic modeling (LCDM) to analyze the RW performance of Chinese EFL students. She found strong correlations among source use, content, and language use, with language use being the most influential. In a recent study by Xu et al. (2025), the effect of independent reading and writing skills and writing strategies on 219 Chinese 11th grade EFL learners' RW performance was investigated. Based on the findings, connecting and evaluating strategies revealed significant positive correlation with IW performance. Furthermore, it was specified that 29.9 % of the variation in IW scores could be explained by independent reading and writing skills, which was underscored by the data gathered from the interviews. Hence, the strategic integration of independent language skills and writing strategies was recommended for the fulfillment of IW tasks. The above review reveals that studies in this area have chiefly indicated the positive effect of content support (Golparvar & Rashidi, 2021; Jung, 2020; Neumann et al., 2019; Plakans, 2008; Plakans & Gebril, 2012; Xu et al., 2025), correlations between reading comprehension ability and synthesis writing (Lee, 2023; Plakans & Gebril, 2012), little considerable difference in IW ability due to variations in proficiency and topic (Weigle & Parker, 2012), and efficacy of instruction on synthesis writing in IW and writing process (Zhang, 2013).

Studies on Multimodal IW

Some other studies have examined the use of multimodal source texts—specifically reading and listening—in IW. Cumming et al. (2005) analyzed essays from ESL learners across three proficiency levels, comparing RW and LW integrated, and independent writing tasks. Results showed that integrated tasks led to shorter texts, broader vocabulary use, and paraphrasing or summarizing of source materials, especially among more proficient writers. However, no major differences were found between RW and LW tasks. Plakans and Gebril (2013) examined TOEFL iBT IW responses, finding that use of reading and listening source materials explained about 50% of score variance. Listening source use and incorporation of key ideas were stronger predictors of performance than reading source use or verbatim copying. Higher-scoring writers avoided direct copying and focused on key ideas. Soleimani and Mahdavi-pour (2014) investigated how proficiency level and task type (RW vs. LW) influenced textual features and employment of source information in Iranian EFL learners' IW. Results showed that both variables significantly affected features like text length and source use. Low-proficiency learners produced shorter texts showed lower grammatical accuracy, and used more plagiarism. As for task type, the RW task appeared to be longer and included more instances of plagiarism than the LW one. Plakans and Gebril (2016) analyzed TOEFL iBT responses to two RLW integrated tasks to examine how discourse features—cohesion, coherence, and organization—related to writing scores. Coherence and organization were found to be significantly associated with scores, while cohesion markers were not. Cheong et al. (2018) examined how reading and listening comprehension influenced IW among first language (L1) Chinese secondary school students in Hong Kong. Participants completed independent reading and listening tasks, followed by an IW task. Results showed that reading comprehension more strongly predicted writing performance than listening, and higher-order cognitive sub-skills (Elaborating, Evaluating, and Creating) significantly correlated with writing outcomes. Abrams (2019) studied intermediate German learners at a U.S. university, comparing writing tasks with and without content support over ten weeks. Tasks using authentic, especially multimodal, source texts improved grammatical and lexical accuracy, fluency, and reduced cognitive load. Homayounzadeh et al. (2019) examined how source text features influenced performance on two TOEFL iBT RLW integrated tasks among Persian candidates. They found significant differences in comprehension, citation use, and source adoption strategies across tasks, suggesting that variations in source features affect IW outcomes. Machili et al. (2019) investigated the effect of a seven-session explicit strategy instruction on IW among Greek EFL learners. The experimental group, trained in selecting, organizing, and connecting source ideas from reading and video texts, outperformed the control group—especially at intermediate and advanced levels—on overall writing, verbatim use, and citation. However, the short duration of the intervention, absence of a delayed posttest, and lack of qualitative data (e.g., interviews) limited understanding of long-term effects and learner challenges. Uludag et al. (2019) further examined multimodal source use by analyzing how number of ideas, content correctness, and modification influenced RLW integrated task scores. While source idea quantity and content correctness predicted performance, the study did not explore why few ideas were used and did not assess the depth of content accuracy. Modification was treated as minor change; thus without analyzing types of linguistic changes, the study provided limited insight into its role. Cai and Yan (2024) examined the effect of Verbatim Source Use (VSU) on linguistic complexity in integrated (multimodal) argumentative essays by Chinese students. Using factor analysis and mixed-effect models, they found that VSU had a weak direct effect and no significant indirect effect on linguistic complexity, compared to proficiency level. More recently, an investigation by Golparvar et al. (2025) highlighted a significant improvement in phrasal and clausal indicators of syntactic complexity in source-based multimodal writing (summary writing and graph writing) over two semesters. The phrasal index also indicated more syntactic complexity. Task type significantly influenced the obtained findings; that is, summary writing manifested higher syntactic complexity values when compared to graph writing. Similar to the previous section, the aforementioned studies also indicate how multimodal content support (Abrams, 2019; Uludag et al., 2019) and strategy instruction (Machili et al., 2019) may positively affect IW performance. Moreover, the studies show how proficiency level of learners may affect features of their IW (Cai & Yan, 2024; Soleimani & Mahdavi-pour, 2014), and how task type (Cumming et al., 2005; Golparvar et al., 2025; Plakans & Gebril, 2013; Soleimani & Mahdavi-pour, 2014) and task features (Cumming et al., 2005; Homayounzadeh et al., 2019; Plakans & Gebril, 2016) may or may not influence different aspects of multimodal integrated tasks.

In summary, the literature review shows that IW tasks have been widely studied, with many focusing on source text use (e.g., Abrams, 2019; Cai and Yan, 2024; Cumming et al., 2005; Golparvar & Rashidi, 2021; Jung, 2020; Plakans & Gebril, 2013). However, just a few studies have examined the effects of instructional interventions (e.g., Machili et al., 2019; Zhang, 2013), in the short run and in limited EFL or English as a Second Language (ESL) contexts. Most research focusing on the employment of both reading and listening source texts has used RLW tasks (e.g., Homayounzadeh et al., 2019; Plakans & Gebril, 2013, 2016; Uludag et al., 2019) under time constraints (Uludag et al., 2019) with limited use of separate RW and LW tasks (e.g., Cumming et al., 2005; Soleimani & Mahdavi-pour, 2014). Moreover, findings on the relative contribution of reading vs. listening sources remain mixed (Cheong et al., 2018; Plakans & Gebril, 2013). This divergence may be attributed to the characteristics of participants of the studies, the former examining L1 Chinese secondary school students aging 17 but the later investigating L2 learners of English with 47 different L1 backgrounds ranging in age from 14-51. In addition, few studies have employed mixed-methods designs (e.g., Homayounzadeh et al., 2019; Plakans & Gebril, 2012). As Golparvar and Rashidi (2021) noted, research in the Iranian EFL context is also scarce. To bridge the research gaps felt, this study aims to longitudinally examine the effect of source text use instruction on Iranian EFL learners' IW performance across RW and LW tasks. Further, the study intends to explore if there are any differences in terms of integration of source text materials across RW and LW integrated tasks. To address the objectives of the study, the following research questions are proposed:

Research Question One: Does the instructional intervention on the use of source texts in IW tasks (i.e., RW and LW tasks) contribute to the overall quality of the students' IW performance in the short and in the long run?

Research Question Two: Does the effect of the intervention significantly vary across RW and LW integrated tasks?

Research Question Three: Do participants incorporate source text materials in their RW and LW tasks differently? How?

Methodology

Design

This research was an explanatory sequential mixed-methods research (MMR) study, since it involved both quantitative and qualitative data collection and data analyses. It was in the form of a pre-experimental one-group pretest-posttest-delayed posttest design.

Participants

Participants in the study were a group of 50 male and female Iranian sophomore students of English, whose age ranged from 20 to 25, studying at a state university in Iran. The contributors partook in two intact classes of Essay Writing (offered over 16 weeks), but stood for one experimental group. The classes were taught by the same instructor, with over 15 years of experience in teaching writing courses. Both classes received instruction on source use in RW and LW tasks. Students' proficiency level, assessed at the beginning of the experiment by administering Oxford Quick Placement Test (OQPT) Version 1 (Oxford University Press, 2001), was intermediate and results of an independent samples t-test revealed no significant difference between the groups' mean scores ($t(48) = -.296, p > .05$). From among these participants, 18 volunteered to take part in the interview conducted after the immediate posttest.

Instruments and Materials

Five instruments were employed in the study. As mentioned before, OQPT Version 1 (Oxford

University Press, 2001) was administered to guarantee the homogeneity of participants in the two groups. Three other instruments were the pretest and the immediate and delayed posttests administered before, right after, and three months after the treatment, respectively. Each of these tests consisted of two IW tasks (i.e., one RW and the other LW) assigned to the two classes with different order to reduce the effects of order on the obtained findings. It is worth mentioning that the pretest, the immediate and delayed posttests as well as IW tasks used for practice, described below, were carefully chosen from among sample IW tasks in Test of English as a Foreign Language (TOEFL) to ensure similarity of the source texts in the tests and in the instructional material in terms of difficulty level. Indeed, to make selections, we initially took into account criteria regarded by earlier researchers (e.g., Cho et al., 2013; Li, 2014; Yu, 2009) as significant in affecting the quality of learners' written texts including topic, lexical and conceptual overlap, and structural organization. Additionally, the reading and listening source texts were carefully chosen so that they were of the same genre, i.e., argumentative. Participants were asked to do each task in the tests in 20 minutes.

Finally, interviews were conducted with the 18 volunteers after the immediate posttest in order to supplement the data collected through the tests and to verify the findings. Students were interviewed individually, each lasting about 15 minutes. Interview questions comprised five items (see Appendix A), which were initially formed based on the objectives of the study, further checked by two experts in the field and finally revised based on their comments.

Materials for the study consisted of six IW tasks (three RW and three LW tasks), which were similar to the pretest and the immediate and delayed posttests. They were done by the students in six class sessions, each focusing on one type of task. In addition, instructional materials, including PowerPoint slides and IW exercises were used to instruct the learners how to employ source materials appropriately in the tasks by properly quoting and paraphrasing source information and attributing source materials.

Scoring Rubric

The rubric employed in this study was an analytical IW Scoring Rubric (Yang & Plakans, 2012) which deals with four major writing components (i.e., content, organization, language use and verbatim source use). Each component of the rubric had six rating levels from 0 to 5 and descriptors of the writing proficiency for each particular level. The sum of the scores on these components (4*0-5) made up the total scores (0-20) for participants' IW.

Coding Scheme

Shi's (2004) coding framework was employed to codify verbatim source use systematically. The coding scheme focuses on the two aspects of textual borrowing, including "degree of modification of the source text and presence or absence of attribution to the source" (Weigle & Parker, 2012, p. 121). In other words, the scheme addresses whether the included source materials are referenced (R) or included with no attribution (N). Furthermore, it specifies whether the incorporated texts are exactly copied (C), slightly modified (M) or reformulated (R). Thus there are seven categories of textual borrowing: the ones exactly copied with no reference to the author (NC), slightly modified with no reference to the author (NM), reformulated with no reference to the author (NR), exactly copied (but without quotation marks) with reference to the author (RC), slightly modified with reference to the author (RM), reformulated with reference to the author (RR), and the ones which included direct quotations (including quotation marks) with reference (QC). It is worth mentioning that, after counseling with two expert raters, the first researcher, who coded the essays, added another category to the above-mentioned codes (i.e., Other), including strings of words which could not be laid under the specified categories.

Data Collection Procedures

After the homogeneity of participants in the two classes was ensured, in line with the course objectives, students were first taught important points about essay writing and practiced writing independent

expository essays in the first eight sessions. Actually, data collection started in the ninth session and the study was mainly conducted during the remaining eight sessions of the regular classes (one at the beginning and one at the end were devoted to the administration of the pretest and immediate posttest and six sessions to the treatment). Both classes received identical intervention on how to employ source materials appropriately in IW tasks. Although they were assigned similar RW and LW tasks and the same pretest and immediate and delayed posttests, the order of assignment was counterbalanced. After students’ writings were collected at the end of each session, the first researcher evaluated the written samples based on the rubric (Yang & Plakans, 2012). She also employed Shi’s (2004) modified coding scheme in order to evaluate the way participants made use of source materials in their writing tasks. Supportive feedback, especially concerning source use, was also provided in the written texts to help students revise their essays. Additionally, at the beginning of each session of the instructional intervention, both groups received general feedback regarding their source integration problems. The plan for data collection procedure in the two classes is summarized in Table 1 and sample instructional material, practice exercises and feedback on students’ writing are illustrated in Appendix B.

Table 1
Plan for Data Collection Procedure

Time	Tasks done
Session 1	Administering pretest (RW then LW in Class 1 and LW then RW in Class 2)
Session 2	Focusing on general important points about IW tasks, Assigning sample IW task 1 (LW in Class 1 and RW in Class 2)
Session 3	Focusing on in-text citations, Assigning sample IW task 2 (RW in Class 1 and LW in Class 2)
Session 4	Focusing on how to evaluate IW tasks and how to make use of quotations, Assigning sample IW task 3 (LW in Class 1 and RW in Class 2)
Session 5	Focusing on how to paraphrase, Assigning sample IW task 4 (RW in Class 1 and LW in Class 2)
Session 6	Focusing on how to paraphrase, Assigning sample IW task 5 (LW in Class 1 and RW in Class 2)
Session 7	Introduction of the analytic IW scale employed in the study, Assigning sample IW task 6 (RW in Class 1 and LW in Class 2)
Session 8	Administering immediate posttest (LW then RW in Class 1 and RW then LW in Class 2)
A week later	Interview
Three months later	Administering delayed posttest

To ensure validity of scoring, 10 percent of the first two practice tasks were randomly chosen and evaluated by an experienced rater as well as the first researcher independently; then, their assessments were closely examined and disagreements were resolved before the first researcher continued to assess the following tasks. As for inter-rater reliability of the scores, the experienced rater and the first researcher scored the tasks independently and then correlation between the two sets of scores was estimated. Correlation indexes obtained were high for the pretest ($r=.94, p<.05$), and for the immediate ($r=.96, p<.05$) and delayed ($r=.97, p<.05$) posttests.

Moreover, the first researcher specified and codified T-units in the immediate posttests based on the modified version of Shi’s (2004) coding scheme. To ensure validity of codification, 10 percent of the immediate posttests were randomly chosen and codified by a PhD candidate who was familiar with IW tasks and trained by an expert on how to use the scheme. The inter-coder correlation obtained was relatively high ($r=.89, p<.05$).

At the end of the treatment, 18 volunteers from the two classes were interviewed. Their answers to the questions were recorded, transcribed and coded for analysis. Three months later, the delayed posttest was administered to the participants in a reverse order to the immediate posttest.

Data Analysis Procedures

The data were analyzed quantitatively (using SPSS 25) and qualitatively. As for quantitative analysis,

descriptive statistics of students’ total IW scores and their scores on individual components of the rubric as well as the number of source-based and total T-units (i.e., source-based T-units plus T-units added by the writer) in the RW and LW immediate posttests, and the frequency and percentages of the interviewees’ answers to the interview questions were calculated. To see whether participants’ writing performance significantly improved from the pretest to the immediate posttest, and to explore the stability of any observed differences over time (delayed posttest), two one-way repeated measures ANOVA tests were run, one for RW and the other for LW tests. Besides, to identify any significant difference in the improvement across the two types of IW tests, first of all, the participants’ gain scores from the pretest to the immediate posttest on each test type were computed and then a paired-samples t-test was run for the computed gain scores.

To focus on any differences in the way participants employed source materials in their IW tasks, T-units in the immediate posttest were scrutinized as described below. Since the number of T-units in the RW source text (N=20) was different from that of LW (N=17), the number of source-based and total T-units in RW and LW tests was converted into rates/proportions to make the figures comparable. To compare rates of source-based and total T-units in participants’ RW vs. LW immediate posttests statistically, two paired-samples t-tests were conducted. Additionally, rates of different borrowing patterns identified based on Shi’s (2004) modified coding scheme were compared across the two task types using Wilcoxon Signed Ranks tests. It is noteworthy that whenever multiple comparisons were made in the analysis of quantitative data, the p-value was adjusted for the number of comparisons in order to avoid Type I error.

Moreover, to analyze interview data, the first researcher initially transcribed the data pertaining to each participant (P), who was randomly identified by a number from 1 to 18 to keep the participant anonymous. Then she made use of the software NVivo to codify the data (through open, axial and selective coding). In other words, thematic analysis was done inductively by assigning codes to the relevant segments in the raw interview data and then refined thematic structures were developed. Further, to assess inter-coder reliability, Cohen’s Kappa was calculated, revealing satisfactory agreement (k=.76). Finally, frequencies and percentages of recurring themes pertaining to research purposes were obtained.

Results

The Impact of Instructional Intervention on Participants’ IW Performance

To answer the first research question addressing the contribution of instructional intervention to participants’ IW in the short and in the long run, participants’ pretest, immediate posttest and delayed posttest mean scores on each type of task were compared. For this purpose, descriptive statistics were first obtained for the total scores on RW pretest, immediate and delayed posttests.

Table 2
Descriptive Statistics for the Pretest, Immediate and Delayed Posttest Scores of RW Tasks

Test	Mean	Std. Deviation	N
RW Pretest	14.3800	1.58938	50
RW Immediate posttest	17.3800	1.42700	50
RW Delayed posttest	17.7600	1.13497	50

As Table 2 exhibits, participants’ mean score on RW delayed posttest (M=17.76) is higher than the one on RW immediate posttest (M=17.38), and both are higher than the mean score on RW pretest (M=14.38). To detect any significant differences, a one-way repeated measures ANOVA was conducted. The test’s assumptions such as independence of observations, normal distribution,

homogeneity of the variances and Mauchly's Test of Sphericity ($\text{sig.} > .05$) were examined, resulting in no violations of the assumptions. The results of the test can be seen in Table 3.

Table 3

Results of Repeated Measures ANOVA for RW Pretest, Immediate Posttest and Delayed Posttest

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	.914	255.19 ^a	2.000	48.000	.000	.914
	Wilks' Lambda	.086	255.19 ^a	2.000	48.000	.000	.914
	Hotelling's Trace	10.630	255.19 ^a	2.000	48.000	.000	.914
	Roy's Largest Root	10.630	255.19 ^a	2.000	48.000	.000	.914

The table shows that there is a statistically significant effect of time for RW tasks (Wilks' Lambda=.086, $F(2, 48)=255.19$, $p < .001$), meaning that participants' mean scores on RW tasks significantly changed across the three time periods (i.e., before and following the intervention, and three months after the intervention). This is also evident in the large effect size obtained (eta squared=.914). In order to reveal where exactly the differences lay, pairwise comparisons were conducted, the results of which are shown in Table 4.

Table 4

Pairwise Comparisons of Mean Scores on RW Pretest(1), Immediate Posttest(2) and Delayed Posttest(3)

(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	-3.000*	.181	.000	-3.448	-2.552
	3	-3.380*	.148	.000	-3.747	-3.013
2	1	3.000*	.181	.000	2.552	3.448
	3	-.380*	.121	.008	-.680	-.080
3	1	3.380*	.148	.000	3.013	3.747
	2	.380*	.121	.008	.080	.680

As illustrated in Table 4, the difference between each pair of time points is significant ($p < .01$), meaning that participants' mean score on RW pretest is significantly lower than their mean score on RW immediate posttest as well as the one they obtained on RW delayed posttest. Moreover, their mean score on RW immediate posttest is significantly lower than theirs on RW delayed posttest. Next, to make a similar comparison for LW test, first descriptive statistics of the scores were obtained.

Table 5

Descriptive Statistics for the Pretest, Immediate Posttest and Delayed Posttest Scores of LW Tasks

Test	Mean	Std. Deviation	N
LW Pretest	13.6600	2.43788	50
LW Immediate posttest	16.7000	1.63195	50
LW Delayed posttest	17.4400	1.43086	50

According to Table 5, participants' mean score on LW delayed posttest ($M=17.44$) is higher than theirs on LW immediate posttest ($M=16.70$), and both are higher than the mean score on LW pretest ($M=13.66$). A one-way repeated measures ANOVA test was conducted to see if the differences were significant. Similarly, assumptions of repeated measures ANOVA were checked and no violations were observed. Table 6 displays the results.

Table 6

Results of Repeated Measures ANOVA for LW Pretest, Immediate Posttest and Delayed Posttest

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	.834	120.691 ^a	2.000	48.000	.000	.834
	Wilks' Lambda	.166	120.691 ^a	2.000	48.000	.000	.834
	Hotelling's Trace	5.029	120.691 ^a	2.000	48.000	.000	.834
	Roy's Largest Root	5.029	120.691 ^a	2.000	48.000	.000	.834

The table shows a statistically significant effect of time for LW tasks (Wilks' Lambda=.166, $F(2, 48)=120.69$, $p<.001$), meaning that participants' mean scores on LW tests significantly varied across the three time periods (i.e., before the intervention, following the intervention, three months after the intervention). The obtained large effect size (eta squared=.834) also highlights this issue. In order to detect which scores significantly differed, pairwise comparisons were made.

Table 7

Pairwise Comparisons of Mean Scores on LW Pretest(1), Immediate Posttest(2) and Delayed Posttest(3)

(I) time (J) time		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	-3.040*	.247	.000	-3.653	-2.427
	3	-3.780*	.245	.000	-4.386	-3.174
2	1	3.040*	.247	.000	2.427	3.653
	3	-.740*	.121	.000	-1.039	-.441
3	1	3.780*	.245	.000	3.174	4.386
	2	.740*	.121	.000	.441	1.039

Table 7 shows that the difference between each pair of time points is significant ($p<.001$). In other words, participants' mean score on LW pretest is significantly lower than their mean score on LW immediate posttest as well as the one they achieved on LW delayed posttest. Additionally, their mean on the LW immediate posttest is lower than their delayed posttest mean score.

In sum, results presented above indicate that the intervention had a positive effect on participants' performance on both task types, RW and LW, in the short and in the long run. The efficacy of instructional intervention in improving participants' IW ability in general was further investigated by analyzing the data gathered through interviews.

Interviewees generally approved of the intervention efficiency. They believed that instruction of the strategies through PowerPoint slides and the instructor's explanations (100%), sample IW tasks they did in the class (88.88%), revision of their tasks (72.22%), introduction of the scoring rubric (72.22%) and the feedback they received (61.11%) contributed to the betterment of their IW performance. The following quotations illustrate the above points, respectively.

I believe the information provided via PowerPoint slides and the professor's explanations in the class were very informative and helpful since the material was taught through examples and short tasks and it helped us learn each point in the best way possible. The good thing was that the PowerPoint slides were available all the time and it let us review the essential points before each writing task. (P2)

As we were expected to do a writing task in every single session, we could put what we had learned in that session into practice before we started forgetting the instruction. In my view, it was particularly effective because it provided us with the opportunity to see how different points work in action. As a result, we could internalize what we had learned. (P13)

Each time I received feedback from the instructor, I learnt more things and became much more aware of my problems and mistakes and that led me to become more vigilant not to commit those mistakes again. These feedbacks truly led to my progress in the field of IW because if you compare my first assignment with my last one you will come across a big difference between the two. (P15)

When writing the revision, I got to know where and why I had such mistakes in writing. It gave me the opportunity to learn from my mistakes and know how to substitute my incorrect choices of words with the accurate ones. (P11)

The scoring rubric gave us a clearer understanding of what was important in the tasks and what proportion of our attention had to be paid to each part. In addition, paying attention to the scores we received for last tasks were more effective in evaluating our progress than just seeing our errors. (P13)

Besides, 50% of the participants believed that in addition to the items mentioned above, other factors such as class discussion, examples presented via slides, oral explanations of the instructor about the students' mistakes, the instructor's being approachable, understanding and supportive and the class positive learning environment played a role in the improvement of their overall IW ability.

Comparative Contribution of Instructional Intervention to Participants' IW across Task Types

To answer the second research question concerning possible differences in the improvement across RW and LW tasks, participants' gain scores from the pretest to the immediate posttest on the two types of tasks were calculated to make any improvement gained across the two types of tasks statistically comparable. Results revealed that participants' mean gain score on the RW task ($M=3.00$) was very close to that on the LW task ($M=3.04$). To inspect any significant difference, a paired-samples t-test was run. Inspection of the assumptions of the test indicated no violations. Results showed no statistically significant difference between participants' mean gain scores on the two types of tasks ($t(49)=-.15$, $p>.05$) with small effect size of .02, meaning that the instructional intervention made similar contribution to participants' improvement in RW and LW tasks.

To triangulate the above findings, the data obtained from the interviews were scrutinized. Half of the students (50%) believed that their RW ability improved more during the treatment; however, most of these students (about 80%) admitted that the difference was not due to the instructional intervention since it did not have much influence on strengthening one type of task more than the other. Nevertheless, from their viewpoints, any difference could be attributed to their own superior ability in employing written source materials or the challenging nature of the aural materials. The following quotes illustrate their opinions.

I believe my RW IW ability has improved more, not due to the way of teaching or the methods used in the class, but due to my own personal ability to remember written materials much easier than oral ones. Hence, I need more practice for improving my LW IW ability. (P6)

I believe there was a difference between these two. LW tasks were more challenging because we had to listen to different accents and simultaneously take notes and focus on the organization of our piece of writing. (P17)

Among the interviewees, just 5.5% claimed that their LW ability improved more due to the treatment since their RW ability was good enough right from the start.

I think I improved in my listening integrated tasks better. It was easier for me to do the reading tasks from the start but my listening tasks needed so much improvement; by doing the tasks more and more, I was able to learn how to do and manage the listening tasks better. (P9)

Comparable improvement of the two task types because of the instructional intervention was acknowledged by 44.5% of the interviewees. Based on their viewpoints, strategies taught by the instructor regarding the employment of source materials in their IW tasks could be employed in both tasks rather similarly; the only difference was that the application of such strategies in LW tasks was harder because of the challenging and fleeting nature of listening materials.

In summary, results of the analysis of the interview data were in line with those of the analysis of gain scores on the two types of tasks in that neither indicated any difference in the intervention efficacy across the two task types.

Comparative Incorporation of Source Materials across Task Types

As for the third research question focusing on any differences in source text use patterns participants employed in RW versus LW tasks, the RW and LW immediate posttests were analyzed. Since the number of RW source text T-units (N=20) was different from that of LW (N=17), the rates/proportions of RW and LW source-based and total T-units in the immediate posttests were calculated and compared. Table 8 shows the results.

Table 8
Descriptive Statistics for Rates of RW and LW Source-based (SB) and Total T-units

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	RW SB T-units	70.5000	50	15.81945	2.23721
	LW SB T-units	63.0536	50	14.16564	2.00332
Pair 2	RW Total T-units	72.5000	50	15.55930	2.20042
	LW Total T-units	67.1726	50	13.13143	1.85706

The mean rates of RW immediate posttest source-based T-units (M=70.50) and total T-units (M=72.50) are higher than those of LW source-based T-units (M=63.05) and total T-units (M=67.17), meaning that participants used source materials in their RW tasks more than their LW tasks and also their RW tasks were longer than their LW tasks. To investigate any significant difference between mean rates thus obtained, two paired-samples t-tests were conducted. Prior to the conduction of the tests, assumptions such as independence of observations, normal distribution, homogeneity of the variances and lack of significant outliers were checked; no violations were observed. The results of the tests are displayed in Table 9.

Table 9
Paired Samples Tests for Rates of RW and LW Source-based (SB) and Total T-Units

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig (2-tailed)
Pair 1	RW SB T-units	7.44640%	14.12487	1.99756	3.43216	11.4606	3.728	49	.001
	LW SB T-units					4			
Pair 2	RW Total T-units	5.32740%	13.45600	1.90297	1.50325	9.15155	2.800	49	.007
	LW Total T-units								

According to the table, there is a statistically significant difference between the mean rates of RW and LW source-based (t(49)=3.72, *p*<.025) and total T-units (t(49)=2.80, *p*<.025), suggesting that the number of source-based T-units as well as the length of RW tasks was greater than those of LW tasks. The eta squared statistics calculated for the first (0.22) and the second pair (0.13) indicate a large effect size, meaning that the effect of medium of source text was substantial.

Then based on the modified coding scheme, participants' rates of employment of source text information borrowing patterns across the two tasks were compared. Table 10 displays the results.

Table 10

Descriptive Statistics for Rates of Source Text Borrowing Patterns across RW and LW Immediate Posttests

	NC	NM	NR	RC	RM	RR	QC	Other
RW	3.7000	12.8000	32.4000	1.6000	2.7000	6.2000	8.8000	2.3000
LW	5.0572	19.4066	19.8758	1.1760	2.9400	1.7640	8.1148	4.4688

The table shows that in both tasks NR, i.e., reformulated T-units with no reference to the author, (RW=32.4; LW= 19.8), NM, i.e., slightly modified T-units with no reference to the author, (RW=12.8; LW= 19.4) and QC, i.e., direct quotations with reference to the author, (RW=32.4; LW= 19.8) rank the highest whereas RC, i.e., exactly copied T-units with reference to the author, ranks the lowest (RW=1.6; LW= 1.17). As the comparison reveals, in the RW task, participants made use of the reformulated T-units with no reference to the author (NR=32.40), the reformulated T-units with reference to the author (RR=6.20), direct quotation (QC=8.80), and exactly copied T-units with reference to the author (RC=1.60) more than the same categories in the LW task, i.e., (NR=19.87), (RR=1.76), (QC=8.11), and (RC=1.17), respectively. Nevertheless, in the LW task, they made use of the slightly modified T-units with no reference to the author (NM=19.40), instances which could not be embedded in any of the categories specified in Shi's (2004) coding scheme (Other=4.46%), the exactly copied T-units with no reference to the author (NC=5.05), and the slightly modified T-units with reference to the author (RM=2.94) more than the same categories in the RW immediate posttests, i.e., NM=12.80, Other=2.30, NC=3.70, and RM=2.70, respectively.

Since Shapiro-Wilk tests showed violations of normality of distribution of the data in the above table (all p-values<05), to see if the differences in source text borrowing patterns across the tasks were significant, Wilcoxon Signed Ranks tests were conducted. Table 11 exhibits the results. It is worth emphasizing that in interpreting the results, adjusted significance value ($p=.05/8=.006$) was used to avoid Type I error.

Table 11

Results of Wilcoxon Signed Ranks Tests for Participants' Rates of Source Text Borrowing Patterns across RW and LW Immediate Posttests

	Test Statistics ^a							
	NC	NM	NR	RC	RM	RR	QC	Other
Z	1.425 ^c	6.096 ^c	3.947 ^b	.667 ^b	.431 ^c	4.118 ^b	.222 ^b	3.699 ^c
Asymp. Sig. (2- tailed)	.154	.000	.000	.505	.667	.000	.824	.000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks

c. Based on negative ranks

Results revealed statistically significant differences between RW and LW tasks in the use of NR and RR, favoring RW tasks, but NM and Other, favoring LW tasks. This means that in RW posttest, participants incorporated significantly more reformulated T-units without and with reference to the author than in LW posttest, with large effect sizes (.55 and .58, respectively). On the other hand, in LW posttest, they integrated significantly more slightly modified T-units with no reference to the author and the new category adopted in the present study than in RW posttest, with large effect sizes (.86 and .52, respectively). However, the differences in the other source borrowing patterns across the two tasks were not statistically significant.

To verify the above findings, interviewees' ideas about comparative incorporation of source materials across the two types of tasks were compared. From the viewpoints of the majority (83.33%), the employment of source materials across the two task types was not the same since source use, specifically the use of quotations and paraphrasing, was much easier in RW tasks partially because of the permanence of written material (versus the fleeting and transitory nature of oral material) and partly due to their individual differences in the utilization of different source materials.

Yes, since it was harder to use the exact words mentioned in the audios, it was common that we made more mistakes using quotations in listening tasks than reading tasks. Generally, written texts were easier to paraphrase and quote from. (P10)

I think some of the differences can be attributed to individual differences. For example, since I have a better visual memory compared to my auditory memory, I am less likely to have verbatim source use in reading-to-writing even when I do not get to take notes of a particular part. The differences between listening and reading skills might also drive students to use particular strategies to make up for the existed problems by taking different approaches to use quotation, paraphrasing, etc. leading them to use different techniques of source use.... (P13)

Even some of those who believed that they employed source materials across RW and LW tasks similarly (16.66%) referred to the fact that source employment was easier in the former than the latter.

Actually, there was not a big difference in their usage by me. The only difference was in the way I made notes of these two sources. I mean making notes of reading tasks were much easier because they were right in front of my eyes. (P15)

Discussion

Based on the results presented in Tables 2, 3, 4, 5, 6 and 7 as well those of the analysis of the interview data, it may be concluded that the instructional intervention significantly affected participants' IW performance. These findings support earlier studies (e.g., Cooney et al., 2018; Cumming et al., 2016; Machili et al., 2019; Zhang, 2013) which illuminated the positive role of explicit instruction, including teaching of relevant strategy use and practice, in the improvement of participants' overall source-based writing performance. This also endorses some researchers' (e.g., Cho & Brutt-Griffler, 2015; Corden, 2007; Liu et al., 2019; Zhu, et al., 2025) calling attention to the effectiveness and productiveness of instruction including reading and writing skills in their investigations. Similarly, the results corroborate the findings of the researchers (e.g. Badian, 1999; Bergman, 1999; Berninger, 2000; Morris & Leavey, 2006) who highlighted the positive effect of listening instruction on the improvement of other language skills (i.e., reading, writing or speaking).

As the findings pertaining to comparative contribution of instructional intervention to participants' IW performance across task types exhibited, no statistically significant difference was detected in the participants' RW vs. LW source-based writing, which was also endorsed by those who took part in the interview. It is noteworthy that though half of the interviewees believed that they performed less satisfactorily on LW tasks compared to RW tasks, most of them attributed this issue to the ephemeral and demanding nature of the oral materials and their divergent strengths in dealing with different receptive skills. They admitted that the instructional intervention was fruitful for improving both types of tasks and was not targeted to fortify any specific type of IW tasks. These findings are somewhat consistent with those of Cheong et al.'s (2018) investigation, in which a significant relationship was indicated between both independent reading and listening and IW performance of the participants. However, contrary to their findings that cognitive competence associated with reading comprehension, compared to the one related to listening comprehension, could explain more variance in IW, the present study implied that reading and listening source materials were put at the same level when considering RW and LW improvement as a result of participants' exposure to instructional intervention. The observation of no significant difference in the contribution of instructional intervention to the amelioration of participants' RW vs. LW source-based writing is also somewhat inconsistent with the studies which displayed either the greater contribution of listening source materials

(Plakans & Gebril, 2013) or reading source materials (e.g., Chan & Yamashita, 2022; Payant et al., 2019). The present findings may be justifiable on the grounds that there is a positive correlation between both receptive skills (i.e., reading and listening) and the writing skill (e.g., Aktas et al., 2023; Brito, 2015; Kamal et al., 2021), and that the efficiency of instruction of source integration (i.e., reading or listening materials) into writing has been demonstrated by other researchers (Berninger, 2000; Cho & Brutt-Griffler, 2015; Cooney et al., 2018; Cumming et al., 2016; Liu et al., 2019; Machili et al., 2019; Zhang, 2013) as well. Nevertheless, disparity between the results of the present study and the findings of the aforementioned studies may be accounted for on methodological grounds such as different language backgrounds and various age ranges of the participants as well as the nature of the studies--i.e., Chan and Yamashita's (2022) study being a meta-analysis.

Results presented in Tables 8, 9, 10 and 11 as well as pertinent interview data analysis revealed that there were differences between rates of source-based T-units and total T-units in RW and those in LW immediate posttests. As far as source-use borrowing patterns were concerned, there were similarities and differences. Participants inclined to employ non-referenced reformulated and modified source materials, and direct quotation with reference in both RW and LW tests more than other patterns, respectively and the least employed borrowing pattern in both tasks was referenced exactly copied materials. However, significant differences were found between rates of non-referenced reformulated and referenced reformulated materials in favor of RW and non-referenced modified materials and the new category adopted in the present study in favor of LW. Interview data analysis also revealed most participants' dissimilar employment of source materials across the two types of tasks, particularly in the use of quotation and paraphrasing. All interviewees, even those who believed in their similar employment of source materials across the two task types, acknowledged that source text use was more undemanding in RW tasks compared to LW tasks and they considered the distinctive nature of reading and listening skills as well as their own unequal strengths in the utilization of the two types of source materials as the main reasons.

The current findings indicating most participants' tendency to utilize non-referenced reformulated and modified source materials in their IW tasks to a greater extent compared to other source-text borrowing patterns, such as referenced exactly copied materials, are consistent with Gebril and Palakans' (2009) which revealed that all their participants reported to employ source materials mainly for shaping notions and viewpoints, and that they were impartial about plagiarism and exact copy of source materials. Similarly, the findings lend support to Weigle and Parker (2012) who displayed that most of the employed word strings were non-referenced, mainly appeared in the form of the modified ones, and that referenced source text borrowing patterns were indicated to be the least utilized ones. Furthermore, the outcomes are partially in line with Shi's (2004) revealing that slightly modified paraphrase was one of the most common textual borrowing patterns employed by L1 and L2 students in their opinion and summary tasks. Our finding that non-referenced source borrowing patterns were prominently employed also supports Shi's conclusion that L2 learners mostly tended to borrow materials without citation. Additionally, the result indicating that verbatim source employment ranked the lowest in both tasks in the present research is in line with Cumming et al.'s (2005) research, which announced no significant difference between their participants' RW and LW tasks with respect to the examined features including verbatim employment of source materials.

As for comparison across task types, the superior borrowing of source text materials in RW tasks, compared to LW tasks, is in line with Plakans and Gebril (2013) as well as Kyle's (2020) acknowledgement of lower-proficiency participants' greater employment of word strings from reading materials than from listening source texts. This outcome also advocates Soleimani and Mahdavi-pour's (2014) finding that their participants generally tended to write longer compositions when they were using written source materials than oral source materials.

Our finding (i.e., the lowest inclination of the participants to employ referenced exactly copied materials in both tasks) contrasts with Keck's L2 student writers' substantial use of nearly copied materials in their summaries (2006) but aligns with her later study (2014) showing that the strategy reduced as a result of university experience, which may be supposed to be comparable to the instruction our students received. In addition, the fact that most participants utilized the reformulated and modified

source text borrowing patterns without citing the reference much more than exactly copied source materials with citation contradict Grabe and Zhang (2013a) who reported that summarizing and synthesizing source materials were more troublesome for students with lower-level reading comprehension proficiency and that naïve reading skill caused them to mechanically duplicate information from source texts. Hence, the inconsistency in their findings and ours may be, at least, partially ascribed to the differences in language proficiency of the participants.

Finally, direct quotation with reference, exactly copied source materials with and without reference and slightly modified source information with reference were found to be similarly used across RW and LW tasks. This might suggest that modality (oral vs. written text) did not significantly affect the participants' tendency to use copying strategies and slightly modified materials with reference to the author and that it did not mediate the effect of instructional intervention on learners' tendency to employ these strategies.

Conclusion and Implications

Several conclusions may be drawn based on the outcomes of the research. Nevertheless, given the limitations of the study discussed in detail below, they may only be suggestive and should be interpreted cautiously. The findings illuminated that the instructional intervention, feedback and practice led to significant development of participants' IW performance and their satisfaction with the intervention. This may suggest that, consistent with the assumptions of SAT, straightforward provision of necessary information about strategies of appropriate employment of source materials in IW tasks (declarative knowledge), purposeful practice of synthesis writing, and helpful feedback and scaffolding led to participants' proceduralization (performance on IW tasks) and, to some degree, automatization of IW skills (more precise and less demanding IW performance). Moreover, the contribution of instructional intervention to the participants' RW and LW tasks was revealed not to be statistically significant. Though participants' textual borrowing amount and patterns were clarified to be somewhat divergent across their RW and LW performance, most of the participants appeared to utilize (non-referenced) reformulated and modified source materials as well as (referenced) direct quotation more than other source borrowing patterns and resorted to referenced exact copying least of all in both tasks. This shows the efficiency of instruction in terms of helping participants to successfully utilize source materials by paraphrasing, modifying and properly quoting the source information. However, the finding that despite the instruction, participants still predominantly utilized non-referenced source materials may be attributed to the type of source, which was not multiple-authored and thus participants may have considered the information redundant. Alternatively, the finding may be justified based on Skehan's (2014) limited attentional capacity, which implies a trade-off between integration of source information and attribution to the source. Finally, source employment was considered troublesome by the majority of the participants, especially with respect to their LW performance, which highlights the need for further strategic instruction of source borrowing in IW, particularly in LW, to the learners (Xu et al. 2025).

The findings of the present research could be enlightening for several stakeholders. Primarily, with respect to the satisfactory viewpoints of the participating students, instructional intervention dealing with the strategies required for the fulfilment of such tasks should be focused on in writing classes to facilitate learners' source borrowing and source integration processes. In addition, as Neumann et al. (2019) recommend, instructive materials should target the focus through enhancing the perception of source texts as well as amelioration of linguistic tools required for appropriate integration of source materials in the students' writing. This calls writing teachers' attention to such instructional intervention and developing and offering relevant tasks and scaffolding in their writing classes and also the curriculum designers' and material developers' attention to designing appropriate materials in this regard, particularly in academic contexts.

In addition, since source integration was revealed in the current research to be significantly different across the two types of tasks (i.e., with regard to the amount of source borrowing as well as the utilization of quotation, paraphrasing and reformulation of source ideas), writing instructors should

consider where differences exist and in case the incorporation of source materials is not fulfilled appropriately in each type of task, they are recommended to put the problem at the center of their attention and resolve it. For example, if the students have more difficulty with the employment of source materials in LW tasks, as it was the case in the present investigation, instructors are advised to incorporate greater number of such tasks in their writing classes and provide students with strategies specifically focusing on how to deal with the challenges of the tasks.

Finally, as the results showed, EFL learners may have problem appropriately referencing source materials. Since appropriate source text employment accompanied by attribution to the author is a salient feature of academic writing, and hence must be regarded as a component of writing ability construct (Hyland, 2009), writing instructors and material developers should focus on preparing and cultivating materials including utilization of source texts in writing tasks appropriately.

This research, like any other investigation, suffered from some limitations, the consideration of which might inform interested researchers how to endeavor to yield more generalizable findings. Primarily, the small number of participants employed in the study due to their availability and qualifications warrants caution about the risk of overgeneralizing the results. The next limitation is the lack of a control group, in which participants' development in source-based writing is traced in the absence of instructional intervention. Another shortcoming of the study can be the limited number of tasks (i.e., six) collected for instructional purposes. Had larger number of tasks been utilized in the study, different results might have been obtained. Another drawback is not delving into participants' second revisions of the tasks; as mentioned by some interviewees, second revision might have boosted their source integration in IW. Finally, the scoring rubric utilized in the study was introduced toward the end of the treatment. However, as stated by some interviewees, if they had been informed about the scoring rubric and what was expected of them from the very beginning, they may have made more goal-oriented attempts to achieve the standards.

Acknowledgements

The authors would like to sincerely thank respectful anonymous reviewers who provided their insightful comments and constructive suggestions. They extend their gratitude to the participants of the study for their cooperation.

Declaration of Conflicting Interests

The authors declare no conflicts of interests.

Funding Details

This research was conducted without any external funding.

References

- Abrams, Z. I. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task-complexity. *System*, 81, 110-121. <https://doi.org/10.1016/j.system.2019.01.009>
- Aghaalikhani, M., Nazaribagha, K., & Ahmadi-Azad, S. (2022). A mixed-methods investigation of the effect of scaffolding genre knowledge on EFL graduate learners' academic writing skills: Achievements and perceptions. *Iranian Journal of English for Academic Purposes (IJEAP)*, 11(4), 74-91. https://ijeap.cmu.ac.ir/article_168086.html
- Aktaş, N., Özdemir, E. Ç., & Özgül, B. K. (2023). The relationship between reading comprehension and writing skills of fourth grade students and their motivation to read and write. *International Journal of Educational Research Review*, 8(2), 122-136. <https://doi.org/10.24331/ijere.1202499>

- Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension. *Journal of Learning Disabilities*, 32, 138-148. <https://doi.org/10.1177/002221949903200204>
- Bergman, O. (1999). *Wait for me! Reader control of narration rate in talking books*. International Reading Association. <https://scispace.com/pdf/wait-for-me-reader-control-of-narration-rate-in-talking-2r1mhiu655.pdf>
- Berninger, V. W. (2000). Development of language by hand and its connections to language by ear, mouth and eye. *Topics in Language Disorders*, 20(4), 65-84. <https://doi.org/10.1097/00011363-200020040-00007>
- Boscolo, P., Arfe, B., & Quarisa, M. (2007). Improving the quality of students' academic writing: An intervention study. *Studies in Higher Education*, 32, 419-438. <https://doi.org/10.1080/03075070701476092>
- Brito, J. D. (2015). *The effects of listening comprehension on English language learners' writing performance while taking notes* [Master's thesis, Bridgewater State University]. BSU Master's Theses and Projects. <https://vc.bridgew.edu/theses/23>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Buck, G. (2001). *Assessing listening*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732959>
- Cai, H., & Yan, X. (2024). Examining the direct and indirect impacts of verbatim source use on linguistic complexity in integrated argumentative writing assessment. *Assessing Writing*, 61, Article 100868. <https://doi.org/10.1016/j.asw.2024.100868>
- Chan, S., & Yamashita, J. (2022). Integrated writing and its correlates: A meta-analysis. *Assessing Writing*, 54(3), Article 100662. <https://doi.org/10.1016/j.asw.2022.100662>
- Cheong, C. M., Zhu, X., & Liao, X. (2018). Differences between the relationship of L1 learners' performance in integrated writing with both independent listening and independent reading cognitive skills. *Reading and Writing*, 31, 779-811. <https://doi.org/10.1007/s11145-017-9811-8>
- Cho, H., & Brutt-Griffler, J. (2015). Integrated reading and writing: A case of Korean English language learners. *Reading in a Foreign Language*, 27(2), 242-261. <http://commons.und.edu/tl-fac/1>
- Cho, Y., Rijmen, F., & Novák, J. (2013). Investigating the effect of prompt characteristics on the comparability of TOEFL iBT integrated writing tasks. *Language Testing*, 30(4), 513-534. <https://doi.org/10.1177/0265532213478796>
- Cooney, A., Darcy, E., & Casey, D. (2018). Integrating reading and writing: supporting students' writing from source. *Journal of University Teaching & Learning Practice*, 15(5). <https://doi.org/10.53761/1.15.5.3>
- Corden, R. (2007). Developing reading-writing connections: The impact of explicit instruction of literary devices on the quality of children's narrative writing. *Journal of Research in Childhood Education*, 21(3), 269-289. <https://doi.org/10.1080/02568540709594594>
- Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, 10(1), 1-8. <https://doi.org/10.1080/15434303.2011.622016>
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Difference in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10, 5-43. <https://doi.org/10.1016/j.asw.2005.02.001>
- Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes*, 23, 47-58. <https://doi.org/10.1016/j.jeap.2016.06.002>

- DeKeyser, R. (2020). Skill acquisition theory. In B. Vanpatten, G. D. Keating & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (3rd ed., pp. 83-104). Routledge.
- DeKeyser, R., & Criado, R. (2012). Automatization, skill acquisition, and practice in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0067>
- Gebril, A., & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 7, 47-84. <https://www.researchgate.net/publication/237836160>
- Gebril, A., & Plakans, L. (2013). Toward a transparent construct of reading-to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly*, 10, 9-27. <https://doi.org/10.1080/15434303.2011.642040>
- Golparvar, S. E., & Rashidi, F. (2021). The effect of task complexity on integrated writing performance: The case of multiple-text source-based writing. *System*, 99, Article 102524. <https://doi.org/10.1016/j.system.2021.102524>
- Golparvar, S. E., Casal, J. E., & Abdolhasani, H. (2025). The development of syntactic complexity in integrated writing: A focus on fine-grained measures. *Assessing Writing*, 66, Article 100983. <https://doi.org/10.1016/j.asw.2025.100983>
- Grabe, W., & Zhang, C. (2013a). Reading and writing together: A critical component of English for academic purposes teaching and learning. *TESOL Journal*, 4(1), 9-24. <https://doi.org/10.1002/tesj.65>
- Grabe, W., & Zhang, C. (2013b). Second language reading-writing relations. In A. Horning, & B. Kraemer (Eds.), *Reconnecting reading and writing* (pp. 108-133). Parlor Press. <https://wacclearinghouse.org/docs/books/reconnecting/chapter6.pdf>
- Guo, L., Crossley, S. A., McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18, 218-238. <https://doi.org/10.1016/j.asw.2013.05.002>
- Habibi, H., Saleh, A. H., & Singh, M. K. S. (2015). The effect of reading on improving the writing of EFL students. *Social Sciences and Humanities*, 23(4), 1115-1138. https://www.researchgate.net/publication/285219262_The_effect_of_reading_on_improving_the_writing_of_EFL_students
- Homayounzadeh, M., Saadat, M., Ahmadi, A. (2019). Investigating the effect of source characteristics on task comparability in integrated writing tasks. *Assessing Writing*, 41, 25-46. <https://doi.org/10.1016/j.asw.2019.05.003>
- Hyland, T. A. (2009). Drawing a line in the sand: Identifying the borderzone between self and other in EL1 and EL2 citation practices. *Assessing Writing*, 14(1), 62-74. <https://doi.org/10.1016/j.asw.2009.01.001>
- Jung, J. (2020). Effects of content support on integrated reading-writing task performance and incidental vocabulary learning. *System*, 93, Article 102288. <https://doi.org/10.1016/j.system.2020.102288>
- Kamal, R., Pravitasari, H., Rahmah, A. Y., Febriyanti, S., Ristiyana, I., Adhani, N., & Rosyid, I. (2021). A correlation between students' listening and writing skills: A study at an Islamic boarding school. *Journal Fascho in Education Conference Proceedings*, 2(1), 1-6. <https://doi.org/10.54626/PROCEEDINGS.V2I1.117>
- Karimnia, A., & Safnyyat, M. (2021). Investigating source-text use in academic writing tasks of EFL students through integrated writing assessment. *Iranian Journal of English for Academic Purposes (IJEAP)*, 10(3), 71-87. https://ijeap.cmu.ac.ir/article_137617_7bc74a3dbf870e2efa020cfd0c19a91e.pdf

- 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>
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, 25, 4-22. <https://doi.org/10.1016/j.jslw.2014.05.005>
- Kyle, K. (2020). The relationship between features of source text use and integrated writing quality. *Assessing Writing*, 45, Article 100467. <https://doi.org/10.1016/j.asw.2020.100467>
- Lee, K. (2023). Diagnosing Chinese college-level English as a Foreign Language (EFL) learners' integrated writing capability: A log-linear cognitive diagnostic modeling (LCDM) study. *Assessing Writing*, 57, Article 100730. <https://doi.org/10.1016/j.asw.2023.100730>
- Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), 39-69. <https://doi.org/10.2307/3587974>
- Li, J. (2014). Examining genre effects on test takers' summary writing performance. *Assessing Writing*, 22, 75-90. <https://doi.org/10.1016/j.asw.2014.08.003>
- Lim, H., & Godfroid, A. (2014). Automatization in second language sentence processing: A partial, conceptual replication of Hulstijn, Van Gelderen, and Schoonen's 2009 study. *Applied Psycholinguistics*, 36(5), 1247-1282. <https://doi.org/10.1017/S0142716414000137>
- Linuwih, E. R., & Winardi, Y. K. (2020). The influence of reading habit on students' writing skill. *Anglo-Saxon: Jurnal Ilmiah Program Studi Pendidikan Bahasa Inggris*, 11(1), 37-47. <https://doi.org/10.33373/as.v11i1.2432>
- Liu, H., Brantmeier, C., & Strube, M. (2019). EFL test preparation in China: The multidimensionality of the reading-writing relationship. *Reading in a Foreign Language*, 31(1), 44-61.
- Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In Mayo, M. del P. G., Mangado, M. J. G., & Adrian, M. M. (Eds.), *Contemporary approaches to second language acquisition* (pp. 71-91). John Benjamin's Publishing Company. <https://doi.org/10.1075/aals.9.07ch4>
- Machili, I., Papadopoulou, I., & Kantaridou, Z. (2019). Effect of strategy instruction on EFL students' video-mediated integrated writing performance. *Journal of Second Language Writing*, 48, Article 100708. <https://doi.org/10.1016/j.jslw.2019.100708>
- Morris, T., & Leavey, G. (2006). Promoting phonological awareness in nursery-aged children through a Sure Start Early Listening programme. *International Journal of Early Years Education*, 14(2), 155-168. <https://doi.org/10.1080/09669760600661559>
- Neumann, H., Leu, S., & McDonough, K. (2019). L2 writers' use of outside sources and the related challenges. *Journal of English for Academic Purposes*, 38, 106-120. <https://doi.org/10.1016/j.jeap.2019.02.002>
- Oxford Quick Placement Test (Version 1) (2001). *Oxford University Press and University of Cambridge Local Examinations Syndicate*. Oxford University Press. <https://www.vhs-aschaffenburg.de/fileadmin/vhs-aschaffenburg/PDF/OxfordTest.pdf>
- Pawlak, M. (2011). Instructed acquisition of speaking: Reconciling theory and practice. In M. Pawlak, E. Waniek-Klimczak & J. Majer (Eds.), *Speaking and instructed foreign language acquisition* (pp. 3-23). Multilingual Matters. <https://dokumen.pub/speaking-and-instructed-foreign-language-acquisition-9781847694126.html>
- Payant, C., McDonough, K., Uludag, P., & Lindberg, R. (2019). Predicting integrated writing task performance: Source comprehension, prewriting planning, and individual differences. *Journal*

of English for Academic Purposes, 40, 87-97. <https://doi.org/10.1016/j.jeap.2019.06.001>

- Peregoy, S. F., & Boyle, O. F. (2009). *Reading, writing, and learning in ESL: A resource book for K-12 teachers*. Pearson Education. <https://dokumen.pub/reading-writing-and-learning-in-esl-a-resource-book-for-teaching-k12-multilingual-learners-eighth-edition-9780137535477-9780137535613-0137535473.html>
- Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, 13, 79-150. <https://doi.org/10.1016/j.asw.2008.07.001>
- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, 17, 18-34. <https://doi.org/10.1016/j.asw.2011.09.002>
- 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, 217-230. <https://doi.org/10.1016/j.jslw.2013.02.003>
- Plakans, L., & Gebril, A. (2016). Exploring the relationship of organization and connection with scores in integrated writing assessment. *Assessing Writing*, 31, 98-112. <https://doi.org/10.1016/j.asw.2016.08.005>
- Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication*, 21(2), 171-200. <https://doi.org/10.1177/0741088303262846>
- Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.) *Processing perspectives on task performance* (pp. 211-260). John Benjamins e-Platform. <https://doi.org/10.1075/tblt.5>
- Soleimani, H., & Mahdavi-pour, M. (2014). The effect of variations in integrated writing tasks and proficiency level on features of written discourses generated by Iranian EFL learners. *The Journal of Teaching Language Skills (JTLS)*, 6(2), 131-159. <https://doi.org/10.22099/jtls.2014.2060>
- Uludag, P., Lindberg, R., McDonough, K., & Payant, C. (2019). Exploring L2 writers' source-text use in an integrated writing assessment. *Journal of Second Language Writing*, 46, Article 100670. <https://doi.org/10.1016/j.jslw.2019.100670>
- VanPatten, B. & Benati, A. G. (2010). *Key terms in second language acquisition*. Continuum International Publishing Group. <https://doi.org/10.1111/j.1473-4192.2011.00288.x>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>
- Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing*, 21, 118-133. <https://doi.org/10.1016/j.jslw.2012.03.004>
- Xu, Y., Zeng, M., & Huang, M. (2025). Impacts of independent language skills and writing strategies on integrated writing performance of Chinese secondary EFL students. *Journal of Second Language Writing*, 68, Article 101209. <https://doi.org/10.1016/j.jslw.2025.101209>
- Yang, H., Plakans, L. (2012). Second language writers' strategy use and performance on an integrated reading-listening-writing task. *TESOL Quarterly*, 46(1), 80-103. <https://doi.org/10.1002/tesq.6>
- Yu, G. (2009). The shifting sands in the effects of source text summarizability on summary writing. *Assessing Writing*, 14(2), 116-137. <https://doi.org/10.1016/j.asw.2009.04.002>
- Zhang, C. (2013). Effect of instruction on ESL students' synthesis writing. *Journal of Second Language Writing*, 22, 51-67. <https://doi.org/10.1016/j.jslw.2012.12.001>
- Zhu, X., Sun, Y., Liu, Y., Xu, W., & Cheong, C. M. (2025). Towards a better understanding of integrated writing performance: The influence of literacy strategy use and independent

Appendix A

Interview Questions

1. Please express your opinion about the efficacy of each of the items mentioned in improving your listening/reading-to-write integrated writing ability in general:
 - A) the instruction of strategies provided through the instructor's PowerPoint slides and explanations in the class
 - B) the sample integrated writing tasks you did in the class in order to practice what you were taught
 - C) the feedback you received on each of the sample integrated writing tasks before you did the next task
 - D) your revision of the sample integrated writing tasks according to the feedback you received
 - E) the introduction of the scoring rubric
 - F) anything else which may have helped you improve your integrated writing ability
2. Do you think there was any difference in the improvement of your reading-to-write integrated writing ability vs. your listening-to-write integrated writing ability as a result of the items mentioned above? How?
3. Do you think your source text use (i.e., the way you make use of quotation, paraphrasing, and modification and reformulation of source information) differs across the two types of tasks (i.e. reading-to-write integrated writing vs. listening-to-write integrated writing)? How?

Appendix B

Sample Instructional Material

When do you need an in-text citation?

An in-text citation should be included whenever you **quote** or **paraphrase** a source in your text.

- **Quoting** means including the original author's words directly in your text. Quotes should always be cited and indicated with quotation marks.
 - Example:
 - ✓ *Original reading text:* It's also worth noting that children who are harshly disciplined in this way tend to be hostile, defiant, and aggressive socially.
 - ✓ *Integrated text:* According to the writer, it's noteworthy that overly-strict upbringing may cause children "to be hostile, defiant and aggressive socially".
- **Paraphrasing** means putting information from a source into your own words. In-text citations are just as important here as with quotes, to avoid the impression you're taking credit for someone else's ideas. In a good paraphrase, not only the wording is changed but also the structure is modified. Further information in this regard will be provided in detail later.
 - Example:
 - ✓ *Original listening text:* The researchers found that patrons who heard music with slow tempo tended to remain in the restaurant longer than patrons who heard fast music.
 - ✓ *Integrated text:* As the lecturer stated, researchers observed that slower-paced music, compared to faster tracks, caused customers to stay longer in the restaurant.

Sample Practice Exercises

1. Suppose you are going to integrate the following sentence from a reading passage into your own text. Use direct quotation to integrate it into your writing.
 - ✓ *Original text:* It is now becoming more widely accepted that pet therapy can provide needed emotional support to people of all ages, and that there are physical benefits as well.
 - ✓ *Integrated text:* ?
2. Suppose you are going to integrate the following sentence from a lecture into your own text. Paraphrase and integrate it into your writing.
 - ✓ *Original text:* At the end of six month, it was found that those who had been caring for a dog had a decrease in blood pressure, as compared to the patients who had received traditional treatment only.
 - ✓ *Integrated text:* ?

Sample Feedback on Students' Writing

First example:

- ✓ *Original reading text:* One of the main advantages of competition is that it creates an environment in which students push each other to excel and increase productivity.
- ✓ *Students' integrated text:* As the writer remarks, one of the benefits of competition is that it produces an environment in which learners force each other to outperform and augment productivity.
- ✓ *Feedback provided:* Good, but you have paraphrased the source too closely by just employing synonyms for some words without changing the structure.

Second example:

- ✓ *Original listening text:* The ability to feel empathy for fellow creatures is often considered one mark of intelligence and has been the subject of a number of research studies.
- ✓ *Students' integrated text:* The ability to feel empathy for fellow creatures is most of the time regarded as a sign of intelligence, recently considered as some investigations' topic.
- ✓ *Feedback provided:* You have initially inserted the exact words of the lecturer into your text without using quotation marks and citing the reference, and then paraphrased the source too closely.