The Effect of Ipsative Assessment on EFL Learners' Argumentative Writing Development: Complexity, Accuracy, and Fluency in Focus

Hassti Gholam-Shahbazi [1], Mehrdad Amiri [2]*, Neda Hedayat [3]

ARTICLE INFO	ABSTRACT
	Among alternative types of assessment, Ipsative Assessment (IA), as a comprehensive form of dynamic assessment, has recently attracted a lot of attention among researchers. This study was an attempt to determine the effect of IA on English as a Foreign Language (EFL) learners' argumentative writing components such as complexity, accuracy, and fluency (CAF). To so doing, 76 male and female EFL learners within the age range of 18-42 in a language institute in Tehran, Iran were selected out of 95 intermediate-level learners, based on their performance in a standard Oxford Placement Test (OPT). The selected participants were randomly divided into two groups; the experimental group (N = 40) received IA and the control group (N = 36) received teacher assessment. The study participants went through the process of pretesting, intervention, and post-testing. A measure of one-way Analysis of Covariance (ANCOVA) and a Multivariate Analysis of Covariance (MANCOVA) were run to probe the null-hypotheses. The results indicated that IA strategies had a statistically significant effect on EFL learners overall writing performance. In addition, the experimental group significantly outperformed the control group on the posttests of complexity, accuracy, and fluency, as L2 writing sub skills. The study raised some pedagogical implications for EFL teachers and learners such as using IA in the L2 classroom to enrich L2 writing development of learners and some for EFL syllabus designers to encompass IA in the syllabi they design for institutions and language schools. Also, some suggestions for further research were made.

1. Introduction

Teaching and learning writing in a Foreign Language (FL) have been challenging for both teachers and students for a very long time (Al-Jarf, 2022). According to Kushki et al. (2022), of the four L2 abilities, writing presents the greatest challenge to the FL learner. One possible explanation is the intricate dynamics between the writer and reader, who both work to communicate ideas and thoughts through written words. Teaching L2 writing, then, is not just difficult, but also requires specialized training (Zhang & Hyland, 2023). Writing development, according to Hughes et al. (2019), largely depends on students' views, their planning methods, goal setting, drafting, reviewing, rewriting, and editing. Hence, it is not a product-oriented or linguistic knowledge-oriented activity.

In an EFL context, assessment of students' writings is more popular now than ever (Jiang & Hyland, 2024; Teng et al., 2022). Considering the loyalty of the Iranian EFL educational program to traditional assessment, many critics have questioned its usefulness and authenticity pointing to the way

[1] PhD Candidate in TEFL, Department of English Translation, Faculty of Literature and Humanities, Varamin-Pishva Branch, Islamic Azad University, Varamin, Iran. Email: shahbazihassti@gmail.com, ORCID: 0009-0001-3586-5245

[2] Assistant Professor in Applied Linguistics, Department of English Language Teaching, Farhangian University, Tehran, Iran. Email: amirimehrdad@cfu.ac.ir; https://orcid.org/0000-0003-0737-7655

[3] Assistant Professor, Department of English Translation, Faculty of Literature and Humanities, Varamin-Pishva Branch, Islamic Azad University, Varamin, Iran. Email: Neda.Hedayat@iau.ac.ir; https://orcid.org/0000-0002-3934-6741

L2 learners' knowledge is only evaluated by a single score (Arani et al., 2012; Firoozi et al., 2019) relying on a product-oriented mechanism (Rezai et al., 2021). Besides, traditional assessment provides a limited way for the students to demonstrate what they have learned (Abedi et al., 2019; Rezai et al., 2024). In addition, in the traditional assessment of L2 writing, learners feel pressured, which can cause anxiety and stress among them affecting their performance (Firoozi et al., 2019; Norova & Haydarali, 2021). Students' potential is limited and does allow out-of-the-box thinking; in other words, learners thinking is not creative (Kutbiddinova, 2021).

What is of paramount importance in the modern views on assessing L2 writing; such as Ipsative Assessment (IA), is how much progress a student makes individually and not considering if a statistical norm is met as a criterion which should be achieved for all learners without paying attention to their individual differences (Jiang & Hyland, 2024; Hyland, 2020; Zhang & Hyland, 2023). Unfortunately, in the Iranian EFL context, a problematic issue with the current state of testing and assessment is that only the students with the highest grades and test scores are praised. This signifies the product-based writing and minimizes the process-based and developmental nature of writing which could be put into practice through IA (Ebrahimi & Ebadi, 2024). Not only does IA value a process-oriented progress, but it guarantees a sort of cognitive-based achievement for all learners practicing it (Hughes, 2017). In addition, as a performance-oriented assessment type (Univio & Pérez, 2019), IA encourages dynamicity in learning and continual assessment in the L2 classroom in general, and in L2 writing classroom in particular, especially in an EFL context like Iran where the traditional assessment types are on board yet, discouraging learners and increasing their stress and anxiety.

By comparing past and present activities, IA determines how far along students are in their learning process (Hughes, 2011). Students are encouraged to reflect on their performances and engage in discourse with themselves, their peers, and the instructor through this continuous comparison, which is vital (Hughes, 2014; Malecka et al., 2021). According to Hughes (2014), this method of providing feedback is helpful because it encourages students to self-regulate by concentrating on when they have completed tasks and what they need to do next in their learning process. Students' self-regulation, self-efficacy, and self-assessment abilities are all positively impacted by IA, according to recent research (Boucher et al., 2017; McIntyre, 2017; Tilley & Roach, 2017; Winstanley, 2017; Zhou & Zhang, 2017). According to Hughes (2014), this research shows that students are more likely to take an active role in providing and receiving feedback when they use IA. Evidence from studies on English writing assessment shows that IA encourages students to reflect on their own writing, self-regulate, and take charge of their own learning (Malecka et al., 2021; Univio & Pérez, 2019). According to Malecka et al. (2021), when teachers provide comments on their students' work, it can motivate them to seek feedback, boost their confidence, and ultimately lead to better writing. Moreover, feedback from teachers also plays a crucial role in helping students to identify areas for improvement in their writing. By offering constructive criticism and pointing out strengths, educators can guide students toward developing their writing skills further.

With respect to educational views, many studies have reported on the weakness of Iranian EFL learners in L2 writing (Abadikhah et al., 2018; Abedi et al., 2019; Ebrahimi & Ebadi, 2024; Fathi & Rahimi, 2020; Nezakatgoo, 2011; Nezakatgoo & Fathi, 2019; Rezai et al, 2024). Besides, traditional assessment is mostly used in Iran (Arani et al., 2012; Farhady & Hedayati, 2009; Firoozi et al., 2019; Razmjoo, 2011; Rezai et al., 2021; Shams & Tavakoli, 2014). Speaking, reading, writing, and listening, as the four English language skills, are taught in universities in the educational system, but these skills are not assessed with equal degrees of importance (Afshar & Asakereh, 2016; Avanaki & Sadeghi, 2013). Sadly, still nowadays assessment is mostly considered as the teachers' responsibility and in isolation from learning. Furthermore, another problematic issue has to do with motivation. Since writing in a foreign language is complicated and a challenging task, and learners have difficulty generating ideas and setting the ideas into readable text, it can demotivate learners (Alharthi, 2021). Thus, writing in English, when it is not the learners' first language, will create additional problems such as demotivation and anxiety (Almulla & Alamri, 2021). According to Chai et al (2021), motivation is related to academic success and could help academic outcomes. Hence, motivation is an important reference that cannot be ignored in L2 learning and more specifically in L2 writing (Jiao, 2022; Hyland, 2020).

Many studies have explored EFL instructors' inclinations to employ Corrective Feedback (CF) in L2 teaching-learning process (Carpenter et al., 2006; Lee, 2020; Lyster & Ranta, 1997, 2013). Moreover, IA, as a meta-cognitively supported assessment type counts on the dynamicity of learners' development which, in its core concept, relies on a continual assessment of learners' performance by themselves after receiving CF from their teacher (Hughes, 2011, 2017, Hughes et al., 2017), has attracted a lot of attention among scholars in literacy development (Collier, 2018), motivation enhancement (Malecka & Boud, 2021), increasing learners' self-esteem (Farkota & Macneill, 2024), assessing university-level science education (Kukol, 2024), teaching and learning grammar (Chernova & Abbasi, 2022), teaching English for Academic Purposes (EAP) (Albert & Mudure-Iacob, 2021), and L2 writing in the ESL contexts (McIntyre, 2017; Univio & Pérez, 2019). However, not many studies have scrutinized the effect of IF, as a technique in IA, on EFL learners' writing performance (Collier, 2018; Ebrahimi & Ebadi, 2024; Nishizuka, 2022; Univio & Pérez, 2019) and few studies have ever focused on the effect of IA on EFL learners' writing performance measured through CAF.

2. Review of Literature
2.1. Theoretical Framework

The present study was investigated under a metacognitive perspective. It consists of three essential skills: planning, monitoring, and evaluation (Schraw, 1998). *Planning* involves selecting appropriate strategies and providing effective resources to achieve goals, such as making predictions prior to reading. It also involves goal setting, activating previous knowledge, and saving time. *Monitoring* is analyzing the effectiveness of the strategies or plans critically. Lastly, *evaluation* refers to triggering further planning, monitoring, and evaluation through examining the progress being made toward goals (Mahdavi, 2014; Skehan, 1998). Flavell (1979) proposed four subcategories of metacognition: metacognitive knowledge (persons understanding about how they can affect their own cognition), metacognitive experiences (which can be added, deleted, or revised, such as measuring progress or likelihood of success at any one time). The significant and positive relationship between metacognition and writing proficiency among EFL learners has been proven (Jiang & Hyland, 2024; Teng, 2020). Accordingly, it is assumed that improving EFL learners' metacognitive regulation skills is crucial for their academic writing development. However, irrespective of its pros, metacognition suffers some disadvantages: Firstly, it may actively hinder learners' task performance, and then the costs of engaging metacognitive strategies may overcome their benefits. Secondly, metacognitive judgments/feelings such as negative self-evaluation may detract learners from psychological well-being (Norman, 2020).

Enhancing EFL learners' metacognitive regulation skills is essential for fostering their growth in academic writing proficiency. While metacognition plays a pivotal role in guiding students towards self-directed learning and improvement, its drawbacks, such as impeding task execution and amplifying perceived challenges, must not be overlooked. Moreover, the negative impact of metacognitive judgments on learners' psychological well-being underscores the need for a balanced approach to cultivating these skills. By acknowledging these complexities and providing targeted support, educators can empower EFL students to navigate the intricacies of metacognition effectively as they strive for academic success and personal growth.

2.2. Ipsative Assessment (IA)

It is argued that IA is an approach to assessing students' progress and covers an extensive range of assessment and feedback practices (Martínez-Arboleda, 2021). IA values progress rather than achievement and mainly motivates and empowers all learners (Hughes, 2017). 'Ipse' in Latin, meaning 'of the self', IA is similarly known as student referenced assessment (Hughes, 2011, 2017; Irmayanti et al., 2024). Moreover, IA is considered as a system for referencing, which reports learning as an approach, which facilitates learner development (Crosby, 2021). Lately, diverse kinds of alternative assessments have been used in various educational settings: Some alternative assessments are journals, portfolios, writing samples, and interviews (Tsedev & Purev, 2024). Particularly, formative assessment, as an alternative assessment, is a means for further progress of learners' development (Hyland, 2020).

By providing feedback in formative assessment the gap between students' performance and criterion-referenced demand is closed while in IA, students' progress is checked by using their

performances as a criterion (Univio & Pérez, 2019). Moreover, in IA, feedback is based on comparing students' past with their present performance and is connected to long-term improvement, whereas feedback in formative assessment is connected to external standards and norms instead of the process of learning (Hughes, 2011). In addition, since IA deals with personal achievement, it differs from other sorts of assessment. Learners become aware of their improvement since they are assessed through self-referential goals and learners become aware of their achievement and realistic progression over time (Univio & Pérez, 2019). As Hughes (2001, 2017) points out shifting from criteria-driven feedback to IF may encourage assessment for learning. The overarching goal of IA is for students to think critically about their own learning, articulate their understanding of course material, and suggest ways to improve in response to instructor feedback (Hughes et al., 2019). The teacher gives the students constructive criticism by comparing their present and previous work and by highlighting the importance of generalized abilities. The term for this method is *feed-forward* assessment, and it is based not on criteria but on students' individual performance (Hughes, 2011, 2017).

As Hughes (2011) presents, should the learner pay attention to the learning process and self-awareness level, they will benefit from IA manifested through feed-up, feedback, and feed-forward. In *Ipsative feed-up*; students are encouraged to focus on their goals, as there is a switch of focus from process to product. In *Ipsative feedback*; by comparing learners' present work against their past performance, they become aware of their progress, and in *Ipsative feed-forward*, teachers provide feedback only about generic skills. As a means of helping students through life-long learning processes, the emphasis of feedback is beyond surface mistakes (Hattie & Timperley, 2007)

Within the scope of IA, IF provides feedback to a student's work directly, indirectly, orally or through the written form and provides the learner with guidelines to compare their current performance in any skill; (i.e., L2 writing), with their previous performance in the same skill (Martínez-Arboleda, 2021). This process of comparing learners' performances makes them more conscious. Hence, they concentrate more on the feedback they receive from the teacher who plays the role of a scaffold (Hughes, 2011). Besides giving longitudinal feedback at each step, the main purpose of utilizing IA is to give an instrument to learners through which they can overcome their shortcomings, analyze the feedback that they receive, and ask for specific feedback if it is essential (Hughes et al., 2019). It is argued that IA is considered as a systematic approach (Flood & Romm, 2018) to processes of power in learning organization, mentioning that self-assessment which is embedded in the IA creates a loop learning and is a facilitative intervention in the learning process (Flood & Romm, 2018). This is why Mason (2021) considered that learning about noticing makes learners' go through noticing and get benefit from the feedback they receive. In the same vein, Hanna et al. (2014) also stressed that noticing is an integral part of IF and self-assessment. Moreover, the knowledge of alternative assessment (i.e., IA) is considered as a significant part of assessment literacy of EFL teachers (Dorri et al., 2025).

Previous research (Collier, 2018; Hughes, 2011; Univio & Pérez, 2019) proposes that in line with IA pedagogical implementation is divided into three stages; *diagnostic stage* (aiming at determining learners' ability to improve and arrange their ideas into a clear form; e.g., students' may be asked to write a descriptive composition), *input awareness stage* (intending to help learners go throughout genre analysis and learn from the authentic texts), and *process writing and IA* (trying to familiarize learners with IA through the teachers' explanations and practicing process writing). According to Hughes et al. (2011), using a scaffolding approach to teaching in which an expert in the field offers students assistance to help them progress is the bridge connecting IA to IF. In addition, recent research has revealed that employing screencast technology, which records teachers' feedback on students' writing, can promote IA and learners' self-regulation (Ebrahimi & Ebadi, 2024). However, the preferences of less knowledgeable versus more privileged learners in relation to the feedback presented is a crucial factor that cannot be neglected. In this regard, as Hyland (2020) argues, continual self-assessment can promote learners' L2 writing wiseness. Moreover, considering writing as a cognitive construct, Jalilzadeh and Coombe (2023) suggest learner-oriented assessment as a step forward to involve learners in the process of classroom assessment and motivate them to rely on self-assessment. In the same vein, Jalilzadeh et al. (2024) consider teachers' updating of assessment knowledge as a key element in promoting learners' L2 classroom development. Moreover, Amiri et al. (2024) suggest teacher assessment literacy and its continual promotion as a fundamental part of Pedagogical Content Knowledge (PCK) of foreign language teachers playing a significant role in their

professional development. Likewise, IA has been considered effective in developing a new model of PCK relying on expanding peace education among pre-service EFL teachers with an emphasis on L2 writing and reading (Amiri & Azizi, 2024). Also, alternative assessment types including IA and portfolios can reflect learning-oriented assessment (Esfandiari et al., 2025). In such a perspective, in line with the theory of moving towards a personal best manifested through IA in higher education (Hughes, 2011), continual self-assessment, as a cognitive and meta-cognitive process, can help learners think deeply about their own understanding of peace while writing or reading. Moreover, Gandhi (2017) found that IA made elementary learners compete with themselves and could increase their capability awareness.

IA has been less taken into account in the in-service teacher training courses of institutions (Ebrahimi & Ebadi, 2024). Nor has it been considered significant in the testing and assessment courses offered to pre-service teachers in the TEFL teacher education program in the Iranian context (Amiri et al., 2024). Hence, doing research on this type of assessment and informing foreign language teachers of its benefits can be advantageous. Moreover, employing IA in the L2 writing classroom and investigating its impact on the overall L2 writing and writing subskills such as complexity, accuracy, and fluency of learners can shed light on how effectively this type of assessment can be put into practice.

2.3. Argumentative Writing

The capacity to analyze critically and present one's arguments in writing is a talent that will be in high demand in the modern world (Barasa, 2024). In order to help the reader make a well-informed choice about the subject, argumentative writing involves gathering and presenting evidence. The goal of writing argumentation is to persuade the reader to think deeply and logically about the topic argued (Barrot & Gabinete, 2021). According to Nippold et al. (2005), students are expected to take a position and back it up in argumentative writing. The goal is to persuade the reader to see things from their perspective or to do what they say. Whether in writing, speech, or a public presentation, an argument is the author's effort to persuade the audience by sound reasoning. Usually, it is a separate act of communication with clear time or place constraints (Voroboj, 2009). An argument is a fundamental component of argumentative writing; it is created when the writer seeks to persuade a specific audience to accept something by presenting proof or evidence (Murtadho, 2021). The process of writing an argumentative essay encourages students to think critically about a given issue, do their own research, and then construct and assess evidence in support of their position. Taking a stand on a topic and supporting it with evidence from reputable sources is a common assignment in argumentative writing classes (Setyowati et al., 2017). Additionally, Ferretti and De La Paz (2011) argued that students would not be ready to write argumentative essays, if they were not acquainted with its techniques and assessment standards. Moreover, the process of critical thinking is intimately linked with arguments, where an argument is defined as a collection of statements (Indrilla & Ciptaningrum, 2018).

The goal of every piece of argumentative writing should be to persuade the reader to adopt the author's point of view by providing convincing evidence and reasoning. A writer cannot just present some information; they need to present it in a way that makes sense and leads readers to believe what they are reading (Fisher, 2013). The idea that one may improve their argument by engaging in metacognition is widely held (Barrot & Gabinete, 2021; Indrilla & Ciptaningrum, 2018; Setyowati et al., 2017). Writing an argumentative essay is a wonderful way to practice critical and metacognitive thinking. The ability to write an argumentative essay smoothly is correlated with the writers' metacognitive and critical thinking skills (Fisher, 2013). On the other hand, Ong and Zhang (2010) found that there were no statistically significant differences in general writing patterns (like pausing and flowing) between the descriptive and evaluative writing assignments. Their study is a quantitative survey on writing fluency, with the independent variables being the amount of time spent planning, providing supporting ideas and macrostructure, and having access to drafts during revision. Writing fluency is measured by the mean number of words produced per minute and the linguistic complexity of the texts (Alghizzi, 2017). A plausible reason might be that the quality of argumentation concepts in writing is unrelated to writing fluency, which is assessed in words per minute.

In a similar study, Univio and Pérez (2019) examined how Ipsative assessment influences argumentative writing. The data of the study was gathered using five instruments; pre questionnaire, post questionnaire, Ipsative feedback form, learners' assignments and learners' assignments grades.

This study was framed upon action research approach and a qualitative/quantitative method was implemented. The results revealed that by grasping the structural and reflective nature, Ipsative assessment enhanced learners' argumentative writing. Besides, comparing different drafts made learners to reflect on their progress and also raise their self-awareness of it. Lastly, the whole procedure made learners aware of the fact that their generic skills was being developed. In the present study, two standard argumentative writing tests adapted from IELTS with different topics were administered as the pretest and posttest.

To sum up, argumentative writing, which was focused in the present study, required both cognitive and metacognitive strategies for its development (Teng, 2020; Teng & Zhang, 2018; Thomas & Rose, 2018; Yabukoshi, 2021; Yeni et al., 2021). Hence, it was assumed that IA could most likely help learners engross such strategies in their L2 wiring.

2.4. Writing Complexity, Accuracy, and Fluency (CAF)

The first paradigm addressing CFA to measure L2 acquisition levels was proposed by Skehan (1998). CAF has been used as a separate variable in SLA research since the 1990s, when cognitive psychology and psycholinguistics got better (Anderson, 1993; Levelt, 1993). It has been used to measure both the speaking and writing skills of the learner (Guillot, 1999; Hilton, 2008; Housen et al., 2005; Skehan & Foster, 2007; Towell, 2007) (Alghizzi, 2017; Şahin Kızıl, 2023). It is said that *complexity* is "the formal nature or semantic-functional attributes of L2 items (for example, forms, meanings, and form-meaning mappings" (Housen et al., 2012, p. 4). In the present study, complexity was measured by the proportion of clauses to T-units (Plakans et al., 2016). *Accuracy* is defined as "the ability to avoid errors in performance, possibly reflecting higher levels of control in the language" (Ellis, 2019, p. 545). Moreover, *accuracy* is considered the capacity to construct sentences that are grammatically correct. It is often seen as the most basic CAF structure. In the present study, accuracy was computed in accord with the percentage of error-free T-units based on Plakan et al. (2016) derived from Skehan and Foster's (1999) study, arguing that global units (i.e., T-units) are more practical and rational measures of accuracy. On the other hand, *fluency* refers to an individual's general language proficiency (Craven, 2017) and is estimated as "the number of words or structural units a writer is able to include in their writing within a certain amount of time" (Wolf-Quintero et al., 1998, p. 14). In line with Alghizzi (2017) and Plakans et al. (2016) who have used the analytical approach to scoring L2 writing, fluency was based on counting the words in error-free production units or just production units (e.g., T-units and clauses).

Now, when evaluating one's language proficiency, it becomes evident that accuracy and fluency play crucial roles in determining one's linguistic capabilities. Accuracy, the ability to craft grammatically sound sentences, sets the foundation for effective communication (Ellis, 2019). Meanwhile, fluency encompasses the overall language proficiency, ensuring a seamless and natural flow of expression. Additionally, the speed at which one can incorporate words and structures into their writing further highlights their accuracy skills (Ellis, 2019; Plakans et al., 2016). Together, these components provide a comprehensive view of an individual's language proficiency, showcasing not only their grammatical prowess but also their ability to convey thoughts with ease and efficiency (Alghizzi, 2017).

Having in mind the issues presented above, the researchers noticed the necessity to use IA as an alternative approach in assessing L2 writing in order to develop learners' abilities to get benefitted from the feedback presented I the L2 writing classroom and enhance their L2 writing skills by noticing the errors they have committed in their previous performances. With regard to the study purposes and the likely significance of IA in promoting EFL learners' argumentative writing, the researchers raised the following questions:

1. Does Ipsative assessment affect EFL learners' overall writing performance?
2. Does Ipsative assessment affect EFL learners' writing components (i.e., complexity, accuracy, and fluency)?

3. Method
3.1. Participants and Setting

Selected out of a population of 95 EFL learners through convenience sampling, the study participants were 76 intermediate-level female (N = 64) and male (N = 12) individuals at TEFL Research Center in Tehran with an average age of 29.11 years (SD = 1.94). The selected EFL learners were those whose scores fell within one standard deviation above and below the mean on a standard Oxford Placement Test (OPT). They were then assigned randomly to the experimental group (N = 40) receiving IA (abbreviated as IAG) and the control group (abbreviated as CG) receiving teacher assessment (N = 36).

3.2. Instrumentation

Learners' homogeneity was gained through OPT whose reliability had been reported as (α = .91) based on Cronbach's alpha (Berthold, 2011, p. 674). Moreover, the construct validity of the test has been confirmed in different studies (i.e., Motallebzadeh & Nematizadeh, 2011; Wistner et al., 2009). The reliability of OPT was estimated in the present study based on KR-21 as (r = 0.87). In addition, two standard argumentative writing topics selected from among the standard topics for IELTS available at (https://www.ielts-blog.com) were used as the writing pre- and post-tests. The participants' writings were scored analytically using Plakans et al., (2016) writing rubric covering CAF. Two raters scored learners written performances and the results of Pearson correlations indicated that there were significant agreements between the two raters on the pretests of complexity [r (78) = .685], accuracy [r (78) = .679], and fluency [r (78) = .772], and posttests of complexity [r (78) = .752], accuracy [r (78) = .608], and fluency [r (78) = .786], representing a large effect size; p < .05.

3.3. Procedures

Once the subjects were selected and homogenized through the results of OPT, they were asked to take the writing pretest. The treatment period lasted 10 sessions. The semester lasted five weeks, with participants attending courses twice a week in 90-minute sessions. In order to make the class more interesting and to increase the effectiveness of teaching L2 writing classrooms, the researchers divided the participants into smaller classes.

In the experimental group (IAG), students received IA and IF both in an oral and written forms. The learners relied on comparing their current performance in writing against their latest performance (Hughes et al., 2011), thus the learners in this group were asked to keep a portfolio of their writing assignments and to review their previous written performance along with the teacher's feedback. It is important to keep in mind that portfolios differ from IF in the sense that portfolios are "a collection of students' works that indicates their experiences, activities, accomplishments, and progress" (Moya & O'Malley, 1994, p. 2). In the context of writing evaluation, a portfolio is defined as a compilation of written texts utilized for various purposes during a certain duration (Weigle, 2002). In summary, it could be said that in IF, learners' performance was compared against their past performance every session e.g., learners' second writing was compared against their first writing, while learners' third writing was compared against their second writing. In order to implement IA in the L2 writing class, the researchers designed the following instructional steps:

By using a PowerPoint, the teacher (who was one of the researchers) explained the IA concept to the students and familiarized them with IA as an alternative form of assessment, IF as a teaching L2 writing strategy, the study objectives, and the procedure to be followed throughout the semester. Hence, the teacher's role was identified as a scaffolder, designer, planner, and facilitator. Then, the teacher introduced an argumentative writing topic and explained how process-based writing (Hedge, 2005), which is used in the IA approach (Hughes, 2011), can be put into practice: Hence, the five steps of *brainstorming* (recalling prior knowledge or schemata for generating ideas), *outlining* (organizing the gathered ideas in an outline, based on rhetorical structure which helped concentrate on the structure of the text), *drafting* (writing the first version of the argumentative text), *revision* (improving the identified aspects of the written work), and *finalizing* (finishing the revised work) were introduced and practiced in the class (Hughes, 2017).

Once the learners in the IAG became familiar with the five steps of process-based writing and IA, they worked on a specific topic. Every two sessions a new standard argumentative writing chosen from (https://www.ielts-blog.com) was given to the students and in every other session 10 students received feedback, meaning that the teacher read the students' first draft and provided them with direct/indirect CF both orally and in the written form, enriched through focused/unfocused CF and metalinguistic feedback. In *direct feedback*, usually the right form of the incorrect form is given; however, in indirect *feedback,* the teacher shows the error's location by highlighting or underlying it without providing the correct form. In the *focused* form the teacher only concentrates on what has been taught and avoids what has not been taught even if it is incorrect, but in the *unfocused feedback* all errors (grammatical, lexical, or sociolinguistic) are corrected by the teacher (Ellis, 2009; Lee, 2020). *Metalinguistic feedback*, as Ellis proposes, takes either the form of error coding in which the teacher writes some abbreviated codes representing different kinds of errors in the margin to indicate what problems students have (for instance, prep=preposition) or a short grammatical description of the committed errors presented briefly at the end of the text (Ellis, 2009; Zhang & Hyland, 2023).

Students were expected to revise their work on their own and submit it to the teacher again. After receiving the learners' second draft the teacher compared it with the first draft and left comments if needed. Finally, the learners were asked for the last time to revise their papers and submit them to the teacher. As mentioned before, every two sessions a new standard argumentative topic chosen from (https://www.ielts-blog.com) was introduced. The learners were free to search the Internet, read books/articles, and consult others to gain enough information regarding the topic and later follow the five stages of process writing in order to hand it over to the instructor. In addition, this group was tracked using a data set (time series design). Every session the participants' scores were recorded and relying on time series design, as a subgroup of longitudinal research design (Schnell, 2021), the developments of learners in L2 writing were monitored.

The learners in the control group dealt with regular L2 writing procedures through the traditional method, by which students wrote about a standard argumentative topic (chosen from https://www.ielts-blog.com) and relied only on the teacher's feedback. The learners in this group received one final direct CF both orally and in written form for their writing by the teacher. However, if learners in this group had any questions regarding how to make use of a specific structure or word dictions, they were provided with explanations only by the teacher. Moreover, learners in this group relied on the teacher's training, the teacher's CF, and their own ability to write in English. It is noteworthy to mention that the teacher scored the learners' writings in this group in accordance with criterion-referenced assessment in each session. The criteria based on which the learners' writings were assessed encompassed CAF (Plakans et al., 2016). Subsequent to the intervention, the participants in both groups were assigned a new standard writing task. The writing posttest was used to assess the participants' improvement in L2 writing. Besides, the participants' writing CAF, as the sub-categories of writing, were checked.

3.4. Design

The study adopted a quantitative and quasi-experimental design due to the fact that a pretest/ posttest was used and a treatment was incorporated in the study between the two tests, and the study focused on the learning procedure the student experienced (Creswell & Plano Clark, 2023). IA was considered as the independent variable and the three components of CAF and argumentative writing as the dependent variables of the study.

3.5. Data Analysis

The results thereof were put into statistical analysis and reported. The collected data were analyzed through SPSS software, version 26. Descriptive statistics was used to calculate the mean and standard deviation of the OPT which was used for homogenizing participants. The OPT had a reliability

of (r = 0.87). For estimating the inter-rater reliability of writing pretest and posttest measures of Pearson correlations were computed. The results indicated that there were significant agreements between the two raters on the pretest of complexity [r (78) = .685, representing a large effect size, p < .05], pretest of grammatical accuracy [r (78) = .679, representing a large effect size, p < .05], and pretest of fluency [r (78) = .772, representing a large effect size, p < .05]. Furthermore, the results indicated that there were significant agreements between the two raters on the posttest of complexity [r (78) = .752, representing a large effect size, p < .05], posttest of grammatical accuracy [r (78) = .608, representing a large effect size, p < .05], and posttest of fluency [r (78) = .786, representing a large effect size, p < .05]. A one-way ANCOVA was run to compare the experimental and control groups' means on the total posttest of writing performance after controlling for the effect of their entry writing performance ability as measured through the pretest. Likewise, a Multivariate analysis of covariance (MANCOVA) was run to compare the experimental and control groups' means on posttest of writing complexity, grammatical accuracy, and fluency after controlling for the effect of the pretest in order to probe the second research question.

4. Results

First, in line with Field (2024) the assumption of normality of the data was retained and then a one-way ANCOVA was run to compare the IAG and CG's means on the total posttest of writing performance after controlling for the effect of their entry writing performance as measured through the pretest. The assumptions of ANCOVA; linearity between the dependent variable (posttest of writing performance) and covariate (pretest), homogeneity of regression slopes, and homogeneity of variances were retained.: First, ANCOVA assumes that the relationship between the dependent variable (posttest of writing performance) and covariate (pretest) is a linear one. Based on the results displayed in Table 1 (F (1, 75) = 28.95, p < .05, η2 = .676 representing a large effect size) it was claimed that the statistical null-hypothesis as the relationship between the two variable was not a linear one was rejected. In other words; there was a linear relationship between the dependent variable and the covariate.

Table 1
ANOVA Test of Linearity of Relationship between Posttest of Writing Performance and Pretest

			Sum of Squares	Df	Mean Square	F	Sig.
Posttest * Pretest	Between Groups	(Combined)	12480.054	45	277.335	1.574	.086
		Linearity	5103.270	1	5103.270	28.958	.000
		Deviation from Linearity	7376.785	44	167.654	.951	.567
	Within Groups		5991.833	34	176.230		
	Total		18471.887	75			
	Eta-Squared		.676				

Second, one-way ANCOVA assumes that the linear relationship between the dependent variable and the covariate is the same across the groups, i.e., homogeneity of regression slopes. As shown in Table 2, the non-significant interaction between the covariate and the independent variable $F(1, 74) = 2.39$, p > .05, partial η2 = .031, representing a weak effect size, indicated that the assumption of homogeneity of regression slopes was met.

Table 2
Testing Homogeneity of Regression Slopes; Posttest of Writing Performance by Groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	1257.558	1	1257.558	33.288	.000	.305

28

Pretest	3887.157	1	3887.157	102.895	.000	.575
Group * Pretest	90.380	1	90.380	2.392	.126	.031
Error	2871.111	74	37.778			
Total	339403.000	76				

And finally, ANCOVA assumes that the variances of the groups be roughly equal; i.e., homogeneity of variances. As shown in Table 3, the non-significant results of the Levene's test, $F(1, 74) = .581$, p > .05 indicated that the assumption of homogeneity of variances was met.

Table 3

Levene's Test of Equality of Error Variances; Posttest of Writing Performance by Groups

F	df1	df2	Sig.
.581	1	74	.448

As Table 4 displays the experimental group (M = 74.77, SE = .982) had a higher mean than the control group (M = 51.89, SE = .892) on the posttest of writing performance after controlling for the effect of the pretest. This ensures that the IA procedure employed by the experimental group was effective leading to the development of overall argumentative writing ability of EFL learners.

Table 4

Descriptive Statistics; Posttest of Writing Performance by Groups with Pretest

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimental	74.777a	.982	72.822	76.733
Control	51.898a	.982	49.942	53.853

Note. a. Covariates appearing in the model are evaluated at the following values: Pretest = 50.83.

Moreover, as shown in Table 5, the results of one-way ANCOVA [F (1, 74) = 270.59, p < .05, partial η^2 = .778 representing a large effect size] indicated that the experimental group had a significantly higher mean on the posttest of writing performance after controlling for the effect of pretest. Thus, IA had a statistically significant effect on Iranian EFL learners' overall writing performance.

Table 5

Tests of Between-Subjects Effects; Posttest of Writing Performance by Groups with Pretest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pretest	4014.385	1	4014.385	104.376	.000	.575
Group	10407.127	1	10407.127	270.590	.000	.778
Error	2961.490	74	38.461			
Total	339403.000	76				

Multivariate analysis of covariance (MANCOVA) was run to compare the experimental and control groups' means on the posttest of the writing task encompassing CAF after controlling for the effect of pretests in order to probe the second research question. First, data normality was retained, and then the assumptions of MANCOVA; linearity of relationships between variables, homogeneity of regression slopes, and homogeneity of variances were retained: Based on the results displayed in Table 6, below, it can be concluded that the assumption of linearity of relationships between pretests and posttests of components of writing was retained.

Table 6

Test of Linearity of Relationship between Pretest and Posttest of CAF

			Sum of Squares	df	Mean Square	F	Sig.
Post-Complexity * Pre-Complexity	Between Groups	(Combined)	338.827	16	21.177	1.474	.138
		Linearity	193.273	1	193.273	13.452	.001
		Deviation from Linearity	145.553	15	9.704	.675	.798
	Within Groups		905.161	63	14.368		
	Total		1243.988	75			
	Eta Squared		.272				
Post-Accuracy * Pre-Accuracy	Between Groups	(Combined)	528.512	14	37.751	3.419	.000
		Linearity	305.423	1	305.423	27.662	.000
		Deviation from Linearity	223.089	13	17.161	1.554	.122
	Within Groups		717.675	65	11.041		
	Total		1246.187	75			
	Eta Squared		.424				
Post-Fluency * Pre-Fluency l	Between Groups	(Combined)	481.931	16	30.121	2.435	.006
		Linearity	242.718	1	242.718	19.618	.000
		Deviation from Linearity	239.213	15	15.948	1.289	.236
	Within Groups		779.456	63	12.372		
	Total		1261.388	75			
	Eta Squared		.382				

The results in Table 6 indicated that the assumption of linearity was met on writing complexity $F(1,75) = 13.45$, p < .05, $\eta^2 = .272$, representing a large effect size, writing grammatical accuracy $F(1,75) = 27.66$, p < .05, $\eta^2 = .424$, representing a large effect size, and writing fluency $F(1,75) = 19.61$, p < .05, $\eta^2 = .382$, representing a large effect size.

In the same vein, the results displayed in Table 7 indicated that the assumption of homogeneity of regression slopes was retained.

Table 7
Test of Homogeneity of Regression Slopes; Pretest of Writing CAF by Groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	63.010	1	63.010	13.945	.000	.155
Pre-Complexity	184.893	1	184.893	40.920	.000	.350
Group * Pre-Complexity	.410	1	.410	.091	.764	.001
Error	343.396	73	4.518			
Total	25779.000	76				
Group	48.248	1	48.248	13.982	.000	.155
Pre-Accuracy	182.575	1	182.575	52.908	.000	.410
Group * Pre-Accuracy	.512	1	.512	.148	.701	.002
Error	262.260	73	3.451			
Total	21249.000	76				
Group	123.411	1	123.411	50.406	.000	.399

Pre- Fluency	217.656	1	217.656	88.900	.000	.539
Group * Pre-Fluency	8.597	1	8.597	3.511	.065	.044
Error	186.072	73	2.448			
Total	16031.000	76				

As Table 7 above displays, the non-significant interaction between covariates and independent variable indicated that the assumption of homogeneity of regression slopes was retained on the writing task $F(1,74) = .417$, p > .05, partial η2 = .005, representing a weak effect size), complexity $F(1,74) =$ .091, p > .05, partial η2 = .001, representing a weak effect size), accuracy $F(1,76) = .148$, p > .05, partial η2 = .002, representing a weak effect size), and fluency $F(1,74) = 3.51$, p > .05, partial η2 = .044 representing a weak effect size). The assumptions of homogeneity of variances and homogeneity of covariance matrices were also significant at .000 based on Box's Test of Equality of Covariance Matrices. As the sample sizes are almost equal in the two groups of the study, these assumptions can be ignored (Field, 2024; Pallant, 2016; Tabachnick & Fidell, 2014).

As Table 8 displays the experimental group had higher means on three components of writing in the posttest after controlling for the effect of the pretest.

Table 8

Descriptive Statistics; Posttests of Components of Writing by Groups with Pretests

Dependent Variable	Group	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Post-Complexity	Experimental	19.949[a]	.249	19.453	20.444
	Control	15.076[a]	.249	14.581	15.572
Post-Accuracy	Experimental	18.646[a]	.249	18.149	19.143
	Control	12.979[a]	.249	12.482	13.476
Post-Fluency	Experimental	16.532[a]	.251	16.031	17.032
	Control	10.643[a]	.251	10.143	11.144

Note. a. Covariates appearing in the model are evaluated at the following values: Pre-Complexity = 13.51, Pre-Accuracy = 13.39, Pre-Fluency = 11.24.

Table 9 displays the results of MANCOVA. The results [F (3, 71) = 43.72, p < .05, partial $η^2 =$ .782 representing a large effect size] indicated that there was a significant difference between the experimental and control groups' components' means on the posttest of writing after controlling for the effect of the pretests. Thus, IA had a statistically significant effect on Iranian EFL learners' writing components.

Table 9

MANCOVA; Posttests of Components of Writing by Groups with Pretests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.755	54.553	3	71	.000	.755
	Wilks' Lambda	.245	54.553	3	71	.000	.755
	Hotelling's Trace	3.073	54.553	3	71	.000	.755
	Roy's Largest Root	3.073	54.553	3	71	.000	.755
Group	Pillai's Trace	.782	63.723	3	71	.000	.782
	Wilks' Lambda	.218	63.723	3	71	.000	.782

Table 10 displays the results of Between-Subjects Effects. Based on these results, and the descriptive statistics displayed in Table 3. It can be concluded that in all the L2 writing sub-skills (CAF), the IAG (M = 19.94) outperformed the CG (M = 15.07) on the writing complexity posttest [F (1, 71) = 171.64, p < .05, partial $η^2$ = .699]. Also, the IAG (M = 18.64) could outperform CG (M = 12.97) on the writing accuracy posttest [F (1, 71) = 230.87, p < .05, partial $η^2$ = .758]. In addition, the IAG (M =

16.53) could significantly outperform the CG (M = 10.64) on the writing fluency [F (1, 71) = 245.19, p < .05, partial η^2 = .768], representing a large effect size.

Table 10
Tests of Between-Subjects Effects; Posttests of Writing Components by Groups with Pretests

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	Post-Complexity	374.085	1	374.085	171.649	.000	.699
	Post-Accuracy	506.149	1	506.149	230.872	.000	.757
	Post-Fluency	546.331	1	546.331	245.195	.000	.768
Error	Post-Complexity	161.273	73	2.179			
	Post-Accuracy	162.233	73	2.192			
	Post-Fluency	164.883	73	2.228			
Total	Post-Complexity	25779.000	76				
	Post-Accuracy	21249.000	76				
	Post-Fluency	16031.000	76				

5. Discussion

The first research question of the study aimed to determine if EFL students' overall writing performance was influenced by IA. The findings demonstrated that the use of IA significantly improved the overall writing performance of the participants. Evidence for this conclusion comes from research by Univio and Pérez (2019) on how Ipsative essay assessments could help EFL learners reflect on their own learning and develop a sense of self-awareness. This finding is in line with a recent study on AI and L2 writing (Ebrahimi & Ebadi, 2024), revealing that screencast-based IA engaged participants in Ipsative self-assessment. Hence, their self-regulation skills in writing were enhanced and learners felt supported, knowing that their mistakes are part of the learning procedure. Likewise, in the present study, IA was practiced by using cumulative coversheets to evaluate the writing development of students (Winstanley, 2017). This method helped improve student achievement and created a positive atmosphere with lots of interaction. Also, this finding aligns with Zhou and Zhang's (2017) study advocating the use of IA which improved learners' learning strategies, made learners aware of their learning goals in L2 writing, and fostered learners' self-efficacy. Similarly, the present study findings take support from Gandhi's (2017) study on IA, finding that IA made elementary learners compete with themselves and increase their awareness concerning learning capabilities. Moreover, the study finding in this section is in line with the theory of enhancing L2 writing capabilities through IA (Hughes, 2011) which asserts that moving towards a personal best can be manifested through IA in higher education. This finding can also take support from Hughes et al. (2011) on the ground that they proved IA could increase motivation of distance learners. In addition, the positive effect of IA on L2 writing performance of the learner is in line with McIntyre's (2017) study on raising L2 learners' self-efficacy through IA which showed a consistent emphasis on IA and the development of students' abilities in self-assessment throughout the project. In addition, the present study findings can take support from Collier's (2018) study on IA which proved that IA could pave a secure way for the freshmen to be more comfortable in learning new materials. The same has been probably taken place in the present research in which learners who received their training through IA outperformed the control group who did not practice this method. As L2 writing development is in close connection with noticing, the present study findings in terms of the impact of IA on overall writing of EFL learners is in line with another study on the effect of learning about noticing and L2 writing development (Mason, 2021) indicating that the techniques of IA can increase learners' noticing in L2 writing. This result also is in line with a previous study on the impact of IA and encouraging students to reflect on their progress in writing (Tilley & Roach, 2017). The combination of ipsative and praised feedback enhanced learners' progress. In a study by Nishizuka (2022), it became evident that IA, as a part of inquiry learning in high schools, could help EFL learners gradually improve their overall writing skills.

With respect to the second research question of the study, the findings showed that not only had the overall writing of the learners improved, but their L2 writing command in terms of CAF. This

is in line with the studies focusing on the impact of corrective feedback on the L2 writing components such as Uysal (2010) who found that written corrective feedback (WCF) could significantly affect learners' paying attention to both accuracy and fluency in writing. In terms of the impact of WCF on the development of writing components among learners, the results are in line with Villarreal and Gil-Sarratea's (2020) study revealing that the peer assessment which enhances discussions among learners through collaborative writing significantly affects their writing performance. Moreover, as participants of the IA group could improve in the writing components, it could be deciphered that the assessment type presented could help them measure their personal improvement which is one of the main tenets of IA (Crosby, 2021).

6. Conclusion

The present study highlighted the significance of IA in promoting EFL learners' argumentative writing measured by CAF. It can be concluded that the process-oriented assessment embedded in IA empowers learners to improve their L2 writing abilities with respect to a dynamic procedure which not only benefits from different types of feedback provided by the teacher but enhances learners' awareness toward both meaning and form in L2 writing, giving way to the development of self-assessment and writing wiseness. The findings also indicated the positive role of IF and the Ipsative portfolios in intermediate EFL learners' argumentative writing development. Ipsative portfolios not only preserved students' previous work but also provided instructors with an overview of their students' progress and the developmental writing processes they had adhered to. This improved the efficacy of IA procedures as it allowed teachers to provide feedback that was more tailored to each student. The use of IA appears to have promoted students' self-regulation and self-assessment. In addition, it seems to have developed a socially secure environment in which students could experience L2 writing development and could make mistakes as they progressed through their writing journey. However, the teacher's IA could promote the definition of L2 writing assessment providing learners with an ongoing, process-based accumulative sort of development that is mainly learner-oriented. Based on these results, EFL teachers may use Ipsative portfolios and detailed IA to help learners improve their L2 writing. Moreover, teachers can use IA to promote self-assessment in their students and have them talk to their teachers about their work so that they could get their questions answered and concerns addressed. Since IA allows teachers to access students' prior work for assessment, it emphasizes the integration of Ipsative portfolios and Ipsative feedback into IA.

The study findings have important implications for EFL teachers, EFL learners, and curriculum designers. Teachers might be interested in using IA in their classes, as it combines instruction, evaluation, and motivation at the same time. In other words, while concurrently comparing students' performance with their past performance, the teacher helps them and acts as a facilitator, planner, designer, and scaffolder. By seeing how language learners react to IA, the teachers are better able to predict the future of the language learners. Moreover, the component of Ipsative feed-forward can be addressed by learners with the help of teachers; meaning that teachers will provide feedback only about general L2 writing skills and let students explore their own problems and solve them through self-assessment (Hattie & Timperley, 2007).

Besides, based on the results of the current study IA can encourage a learner to engage in meta-cognitive self-assessment through a comparison of his/her current performance with a prior one. Hence, learners will be aware of their progress and address their learning needs (Chapman et al., 2005). In addition, IA allows learners to start with what they know and go beyond it. Accordingly, their linguistic competence develops along with their cognitive competence (Gee, 2003). Learners' self-esteem and confidence are achieved and competition will be reduced in the IA-based L2 writing classroom. Additionally, learners are given a sense of socially-safe freedom for the sake of respecting their identities and achieving positive and social emotions (Rattray, 2018). In addition, argumentative writing is considered a complex higher-order cognitive function that cannot be measured by traditional memory-replay tests, but it requires an alternative assessment which could check if the learner has improved in developing arguments in his/her writing (Malecka & Boud, 2021; Seery et al., 2019). This seems to be possible in IA (Hughes et al., 2019). Lastly, with IA learners go through the process of noticing and get benefit from the feedback they receive (Mason, 2021).

Additionally, curriculum designers can improve the L2 classroom curricula utilized in language institutes, high schools, and universities by paying more attention to IA. In simpler terms, this study puts forward some implicit pedagogical suggestions for EFL instructors to explore novel ways of teaching writing presented in IA and refine their current writing instruction procedures. Since IA proved to be successful in making the EFL learners aware of their mistakes and errors, it helped the learners to be active and responsible for their learning. Moreover, learners' motivation was increased because in every session they were making progress due to the fact that their new performance was constantly being compared with their last performance.

While the findings of this study are promising, it is vital to acknowledge certain limitations. The study took place with a limited sample size (N = 76). The small amounts of differences between the pretest and posttests of all three components (CAF) of writing can be attributed to the small size of the sample and the short period of treatment time, which took only five weeks. Bearing in mind that the development of surface-level and deep-level skills of writings takes place over a long time, the present researchers believe that further studies with larger samples and with longer times of treatment are required before making valid conclusions. In addition, the present study took writing complexity, accuracy, and fluency as the dependent variables. More research is needed to concentrate on other language skills of speaking, reading, and listening comprehension in order to reach more valid conclusions in terms of the effect of IA on the holistic development of learners' foreign language. Future studies may be needed to replicate the findings with other language skills or components. Also, this study concentrated on intermediate EFL learners. Other studies can be conducted at advanced and beginner levels. Moreover, the quantitative measures for computing the CAF components were T-units. It is possible to carry out the study with other quantitative measures. Lastly, this study was conducted with male and female participants but did not address the likely effect of gender on employing IA in the L2 writing classroom. It will be possible to run a new study to determine the effects of IA on the writing ability of exclusively male or female learners.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Declaration of Applying AI
The authors did not use AI to generate the information presented in this article.

References
Abadikhah, S., Aliyan, Z., & Talebi, S. H. (2018). EFL students' attitudes towards self-regulated learning strategies in academic writing. *Issues in Educational Research*, *28*(1), 1–17. http://www.iier.org.au/iier28/abadikhah.pdf

Abedi, P., Namaziandost, E., & Akbari, S. (2019). The impact of flipped classroom instruction on Iranian upper-intermediate EFL learners' writing skill. *English Literature and Language Review*, *5*(9), 164–72. https://doi.org/10.32861/ellr.59.164.172

Afshar, H. S., & Asakereh, A. (2016). Speaking skills problems encountered by Iranian EFL freshmen and seniors from their own and their English instructors' perspectives. *Electronic Journal of Foreign Language Teaching*, *13*(1), 112–130. https://eflt.nus.edu.sg/v13n12016/afshar.pdf.

Albert, C. N., & Mudure-Iacob, I. (2021). Gamifying formative and Ipsative assessment for English for specific purposes students. In *ICERI2021 Proceedings* (pp. 8196–8204). IATED. https://doi.org/10.21125/iceri.2021.1870

Alghizzi, T. M. (2017). *Complexity, accuracy, and fluency (CAF) development in L2 writing: The effects of proficiency level, learning environment, text type, and time among Saudi EFL learners* (PhD Thesis), University College Cork. https://hdl.handle.net/10468/4815

Alharthi, S. (2021). From instructed writing to free-writing: A study of EFL learners. *Sage Open*, *11*(1), 65–63. https://doi.org/10.1177/21582440211007112

Al-Jarf, R. (2022). Role of instructor qualifications, assessment and pedagogical practices in EFL students' grammar and writing proficiency. *Journal of World Englishes and Educational Practices (JWEEP)*, *4*(2), 6–17. https://doi.org/10.32996/jweep.2022.4.2.2

Almulla, M. A., & Alamri, M. M. (2021). Using conceptual mapping for learning to affect students' motivation and academic achievement. *Sustainability*, *13*(7), 40–59. https://doi.org/10.3390/su13074029

Amiri, M., & Azizi. F. (2024). Is PCK in the TEFL teacher education in need of peace education? EFL pre-service teachers' and teacher educators' views in focus. In A. Khodamoradi & M. Ahmadi (Eds.), *The Proceedings of the 5th National Conference on Pedagogical Content Knowledge (PCK) in Teaching English as a Foreign Language (TEFL)* (p. 23). Farhangian University, Arak.

Amiri, M., Nezakatgoo, B., & Jalilzadeh, K. (2024). Developing and validating a pedagogical content knowledge (PCK) scale for the TEFL teacher education program in the Iranian context. *Research in Curriculum Studies*, *4*(2), 1–23. https://doi.org/10.48310/jcdr.1403.3722

Anderson, J. R. (1993). Production systems and the ACT-R theory. In J. Anderson (Ed.)*, Rules of the mind* (pp.17–44). Erlbaum.

Arani, A. M., Kakia, M. L., & Karimi, M. V. (2012). Assessment in education in Iran. *SA-Educ, 9*(2), 1–10. http://www.nwu.ac.za/sites/www.nwu.ac.za/files/files/p-saeduc/New_Folder_1/3_Assessment%20in%20education%20in%20Iran.pdf

Avanaki, H. J., & Sadeghi, B. (2013). English language teaching (ELT) in Iranian universities in brief. *Theory and Practice in Language Studies*, *3*(12), 2296–2307. https://doi.org/10.4304/tpls.3.12.2296-2302

Barasa, D. (2024). Demystifying the discourse: Techniques to effective academic writing. *Journal of Research and Academic Writing, 1*(1), 13–21. https://doi.org/10.58721/jraw.v1i1.571

Barrot, J., & Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *International Review of Applied Linguistics in Language Teaching*, *59*(2), 209–232. https://doi.org/10.1515/iral-2017-0012

Berthold, M. (2011). Reliability of quick placement tests: How much faith can we place on quick paper or internet-based placement tests. *Australian Journal of Teacher Education*, *35*(6), 674–698.

Boucher, M., Dubé, F., & Creech, A. (2017). The effect of video feedback on the self-assessment of a music performance by pre-university level classical guitar students. In G. Hughes (Ed.), *Ipsative Assessment and Personal Learning Gain: Exploring International Case Studies* (pp. 197–220). Palgrave Macmillan.

Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, *28*(2), 209–236. https://doi.org/10.1017/S0272263106060104

Chai, C. S., Lin, P. Y., King, R. B., & Jong, M. S. Y. (2021). Intrinsic motivation and sophisticated epistemic beliefs are promising pathways to science achievement: Evidence from high achieving regions in the east and the west. *Frontiers in Psychology*, *12,* 448–457. https://doi.org/10.3389/fpsyg.2021.581193

Chapman, C., King, R., & King, R. M. (2005). *Differentiated assessment strategies: One tool doesn't fit all*. Corwin Press.

Chernova, S., & Abbasi, M. (2022). Using Ipsative assessment in teaching and learning English grammar through e-portfolios. *Research Notes, 83*, 25–34. https://hdl.handle.net/102.100.100/603093

Collier, C. (2018). *Exploring the impact of an Ipsative literacy development intervention on first year university students' fear of failure in assessment* (Doctoral dissertation), University College London. https://discovery.ucl.ac.uk/id/eprint/10043556

Craven, L. (2017). Measuring language performance: Complexity, accuracy, and fluency measures. In E. Levyne & D. Dangles (Eds.), *Proceedings of the 2017 WEI international academic conference proceedings* (pp. 25 –27). West East Institute. https://www.westeastinstitute.com/wp-content/uploads/2017/02/Laurence-Craven.pdf

Creswell, J. W., & Plano Clark, V. L. (2023). Revisiting mixed methods research designs twenty years later. In C. N. Poth (Ed.), *The Sage Handbook of mixed methods research designs*, *1*(1), 21–36.

Crosby, R. (2021). *Using ipsative assessment to improve feedback quality and the student assessment experience in university computer science* (Doctoral dissertation), Newcastle University. http://hdl.handle.net/10443/5427

Dorri, A., Heidari Tabrizi, H., & Lotfi, A. (2025). The impact of language assessment literacy enhancement (LALE) on Iranian high school EFL students' knowledge of assessment as learning in writing. *International Journal of Language Testing*, *15*(1), 40–53. https://doi.org/10.22034/ijlt.2024.444115.1327

Ebrahimi, B., & Ebadi, S. (2024). Exploring the impact of screencast-based Ipsative assessment on EFL students' writing development. *Language Testing in Asia, 14*(1), 1–26. https://doi.org/10.1186/s40468-024-00315-y

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, *63*(2), 97–107. https://doi.org/10.1093/elt/ccn023

Ellis, R. (2019). Towards a modular language curriculum for using tasks. *Language Teaching Research*, *23*(4), 454–475. https://doi.org/10.1177/1362168818765315

Esfandiari, R., Arefian, M. Hossein, & Lam, R. (2025). Understanding collective and reflective learning-oriented assessment among Iranian EFL preservice teachers in learner-centered language teacher education. *International Journal of Language Testing*, *15*(1), 91–106. https://doi.org/10.22034/ijlt.2024.460026.1347

Farhady, H., & Hedayati, H. (2009). Language assessment policy in Iran. *Annual Review of Applied Linguistics*, *29*, 132–141. https://doi.org/10.1017/S0267190509090114

Farkota, D. R., & Macneill, D. N. (2024). Ipsative assessment: Growing students' knowledge, skills and self-esteem. *Assessment.* https://www.educationtoday.com.au/news-detail/Ipsative-Assessment-6297

Fathi, J., & Rahimi, M. (2020). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: A case of EFL students. *Computer Assisted Language Learning*, *35*(7), 1668–1706. https://doi.org/10.1080/09588221.2020.1825097

Ferretti, R. P., & De La Paz, S. (2011). On the comprehension and production of written texts: Instructional activities that support content area literacy. In R. O'Connor & P. Vadasy (Eds.), *Handbook of reading interventions* (pp. 326–355). Guilford Press.

Field, A. (2024). *Discovering statistics Using IBM SPSS: Statistics for statistics* (6th ed.). Sage.

Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, *9*(1), 1–14. https://doi.org/10.1186/s40468-019-0078-7

Fisher, A. (2013). *Critical thinking: An introduction (2nd ed).* Cambridge University Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911.

Flood, R. L., & Romm, N. R. (2018). A systemic approach to processes of power in learning organizations: Part II–triple loop learning and a facilitative intervention in the "500 schools project". *The Learning Organization*. https://doi.org/10.1108/TLO-11-2017-010

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Gandhi, S. (2017). Compete with yourself (CWY): Maximizing learning gain in schools. In G. Hughes (Ed.), *Ipsative assessment and personal learning gain: Exploring international case studies* (pp. 221–242). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_11

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, *1*(1), 20–30. https://doi.org/10.1145/950566.950595

Guillot, M. N. (1999). *Fluency and its teaching* (Vol. 11). Multilingual Matters.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, *129*(3), 1311–1353. https://doi.org/10.1093/qje/qju015

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

Hedge, T. (2005) *Writing*. Oxford University Press

Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, *36*(2), 153–166. https://doi.org/10.1080/09571730802389983

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.01hou

Housen, A., Pierrard, M., & Van Daele, S. (2005). Rule complexity and the effectiveness of explicit grammar instruction. In *Investigations in instructed second language acquisition* (pp. 235–269). Mouton de Gruyter. https://doi.org/10.1515/9783110197372.2.235

Hughes, G. (2014). *Ipsative assessment: Motivation through marking progress*. Palgrave Macmillan.

Hughes, G. (2011). Towards a personal best: A case for introducing Ipsative assessment in higher education. *Studies in Higher Education*, *36*(3), 353–367. https://doi.org/10.1080/03075079.2010.486859

Hughes, G. (Ed.). (2017). *Ipsative assessment and personal learning gain: Exploring international case studies*. Springer Nature.

Hughes, G., Hawkes, D., & Neumann, T. (2017). Use of digital technology to capture and support student progress across a taught postgraduate programme. In G. Hughes (Ed.), *Ipsative assessment and personal learning gain: Exploring international case studies* (pp. 105–128). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_6

Hughes, G., Okumoto, K., & Wood, E. (2011). *Ipsative assessment and motivation of distance learners. A report on a center for distance education research project*. Leading Education and Social Research. http://cdelondon.wordpress.com/2010/07/28/use-ofipsative-assessment-in-distance-learning/

Hughes, M. D., Regan, K. S., & Evmenova, A. (2019). A computer-based graphic organizer with embedded self-regulated learning strategies to support student writing. *Intervention in School and Clinic*, *55*(1), 13–22. https://doi.org/10.1177/1053451219833026

Hyland, K. (2020). The communication of expertise: Changes in academic writing. In M. Gotti, S. Maci, & M. Sala (Eds.), *Scholarly pathways: Knowledge transfer and knowledge exchange in academia* (pp. 33–56). Peter Lang Group. https://ueaeprints.uea.ac.uk/id/eprint/74298/

Indrilla, N., & Ciptaningrum, D. S. (2018). An approach in teaching writing skills: does it offer a new insight in enhancing students' writing ability. *A Journal on Language and Language Teaching, 21*(2), 124–133. https://doi.org/10.24071/llt.v21i2.1036

Irmayanti, A. P., Remiswal, R., Khadijah, K., & Syamsi, S. (2024). Revealing the practice of norm-referenced assessment and criterion-referenced assessment in Islamic education learning: A case study at SMPN 24 Padang. *At-Ta'lim: Media Informasi Pendidikan Islam*, *23*(1), 78–90. http://dx.doi.org/10.29300/attalim.v23i1.4327

Jalilzadeh, K., & Coombe, C. (2023). Constraints in employing learning-oriented assessment in EFL classrooms: Teachers' perceptions. *Language Testing in Asia*, *13*(1), 7–21. https://doi.org/10.1186/s40468-023-00222-8

Jalilzadeh, K., Dastgoshadeh, A., & Khosravi, R. (2024). The impact of an in-service professional development course on teachers' language assessment knowledge. *Journal of Modern Research in English Language Studies*, *11*(2), 131–153. https://doi.org/10.30479/jmrels.2023.18114.2152

Jiang, F. K., & Hyland, K. (2024). Writing book reviews: Perceptions and experiences of Chinese novice scholars. *Journal of English for Academic Purposes*, *71*, 101409. https://doi.org/10.1016/j.jeap.2024.101409

Jiao, S., Wang, J., Ma, X., You, Z., & Jiang, D. (2022). Motivation and its impact on language achievement: Sustainable development of ethnic minority students' second language learning. *Sustainability*, *14*(13), 7898. https://doi.org/10.3390/su14137898

Kukol, A. (2024). Combination of Ipsative and sociomaterial assessment methodologies within university-level science education. *Higher Education for the Future*, *11*(2), 220–234. https://doi.org/10.1177/23476311241268970

Kushki, A., Rahimi, M., & Davin, K. J. (2022). Dynamic assessment of argumentative writing: Mediating task response. *Assessing Writing*, *52*. https://doi.org/10.1016/j.asw.2022.100606

Kutbiddinova, S. (2021). Advantages and disadvantages of alternative assessment. *Integration of Science, Education and Practice. Scientific-Methodical Journal, 2*(3), 110–113. https://bilig.academiascience.org/index.php/isepsmj/article/view/130

Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, *49*, 100734. https://doi.org/10.1016/j.jslw.2020.100734

Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT press.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66. https://doi.org/10.1017/S0272263197001034

Lyster, R., & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, *35*(1), 167–184. https://doi.org/10.1017/S027226311200071X

Mahdavi, M. (2014). An overview: Metacognition in education. *International Journal of Multidisciplinary and Current Research*, *2*(6), 529–535. Available at http://ijmcr.com

Malecka, B., Ajjawi, R., Boud, D., & Tai, J. (2022). An empirical study of student action from ipsative design of feedback processes. *Assessment & Evaluation in Higher Education*, *47*(5), 801–815. https://doi.org/10.1080/02602938.2021.1968338

Malecka, B., & Boud, D. (2021). Fostering student motivation and engagement with feedback through Ipsative processes. *Teaching in Higher Education*, *2*(6), 1–16. https://doi.org/10.1080/13562517.2021.1928061

Martínez-Arboleda, A. (2021). Ipsative assessment: measuring personal improvement. In T. Beaven & F. Rosell-Aguilar (Eds), *Innovative language pedagogy report* (pp. 77–82). Research-publishing.net. https://doi.org/10.14705/rpnet.2021.50.1239

Mason, J. (2021). Learning about noticing, by, and through, noticing. *ZDM–Mathematics Education*, *53*(1), 231–243. https://doi.org/10.1007/s11858-020-01192-4

McIntyre, K. (2017). Raising self-efficacy through Ipsative assessment and Feuerstein's instrumental enrichment programme. In G. Hughes (Ed.), *Ipsative assessment and personal learning gain* (pp. 85–104). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_5

Motallebzadeh, K., & Nematizadeh, S. (2011). Does gender play a role in the assessment of oral proficiency? *English Language Teaching, 4*(4), 165–172.

Moya, S. S., & O'Malley, J. M. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students, 13*(1), 13–36.

Murtadho, F. (2021). Metacognitive and critical thinking practices in developing EFL students' argumentative writing skills. *Indonesian Journal of Applied Linguistics*, *10*(3), 656–666. https://doi.org/10.17509/ijal.v10i3.31752

Nezakatgoo, B. (2011). The effects of portfolio assessment on writing of EFL students. *English language teaching*, *4*(2), 231–241. https://doi.org/10.5539/elt.v4n2p231

Nezakatgoo, B., & Fathi, J. (2019). Second language writing through blogs: An investigation of learner autonomy. *Iranian Journal of Applied Language Studies*, *11*(2), 165–190. https://dx.doi.org/10.22111/ijals.2019.5448

Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools, 36*(2), 125–138. https://doi.org/10.1044/0161-1461(2005/012)

Nishizuka, K. (2022). Significance and challenges of formative Ipsative assessment in inquiry learning: A case study of writing activities in a "contemporary society" course in a Japanese high school. *Sage Open*, *12*(2), 1–19. https://doi.org/10.1177/21582440221094599

Norman, E. (2020). Why metacognition is not always helpful. *Frontiers in Psychology*, *11*, 1–6. https://doi.org/10.3389/fpsyg.2020.01537

Norova, M. & Haydarali, A. (2021). Advantages and disadvantages of traditional and alternative ways of assessment. *E-conference Globe, 44*, 195–200.

Pallant, J. (2016). *SPSS survival manual* (6th ed.). Allen & Unwin.

Plakans, L., Gebril, A., & Bilki, Z. (2016). Shaping a score: Complexity, accuracy, and fluency in integrated writing performances. *Language Testing, 36*(2), 161–179. https://doi.org/10.1177/0265532216669537

Rattray, J. (2018). Affect and Ipsative approaches as a counter to pedagogic frailty: The guardian of traditional models of student success. *Higher Education Research & Development*, *37*(7), 1489–1500. https://doi.org/10.1080/07294360.2018.1494141

Razmjoo, S. A. (2011). Language proficiency tests in the Iranian context: Do they represent communicative language testing model? *Journal of Pan-Pacific Association of Applied Linguistics*, *15*(2), 85–96. https://eric.ed.gov/?id=EJ979916

Rezai, A., Alibakhshi, G., Farokhipour, S., & Miri, M. (2021). A phenomenographic study on language assessment literacy: Hearing from Iranian university teachers. *Language Testing in Asia*, *11*(1), 1–25. https://doi.org/10.1186/s40468-021-00142-5

Rezai, A., Naserpour, A., & Rahimi, S. (2024). Online peer-dynamic assessment: An approach to boosting Iranian high school students' writing skills: A mixed-methods study. *Interactive Learning Environments, 32*(1), 306–324. https://doi.org/10.1080/10494820.2022.2086575

Şahin Kızıl, A. (2023). Data-driven learning: English as a foreign language writing and complexity, accuracy and fluency measures. *Journal of computer assisted learning*, *39*(4), 1382–1395. https://doi.org/10.1111/jcal.12807

Schnell, C. (2021). Time Series Designs. *The Encyclopedia of Research Methods in Criminology and Criminal Justice*, *1*, 344 –347. https://doi.org/10.1002/9781119111931.ch69

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*(1), 113–125.

Seery, N., Buckley, J., Delahunty, T., & Canty, D. (2019). Integrating learners into the assessment process using adaptive comparative judgement with an Ipsative approach to identifying competence-based gains relative to student ability levels. *International Journal of Technology and Design Education*, *29*(4), 701–715. https://doi.org/10.1007/s10798-018-9468-x

Setyowati, L., Sukmawa, S., & Latief, M. A. (2017). Solving the students' problems in writing argumentative essay through the provision of planning. *Celt: A Journal of Culture, English Language Teaching & Literature, 17*(1), 86102. https://doi.org/10.24167/celt.v17i1.1140

Shams, N., & Tavakoli, M. (2014). The effect of peer, self, and traditional assessment on Iranian EFL learners' L2 reading comprehension. *Journal of Applied Linguistics and Language Research, 1*(1), 29–44.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, *49*(1), 93–120. https://doi.org/10.1111/1467-9922.00071

Skehan, P., & Foster, P. (2007). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In S. Van Dael, A. Housen, F. Kuiken, M. Pierrard, & V. I. (Eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching*, (pp. 207–226). University of Brussels Press.

Teng, F. (2020). The role of metacognitive knowledge and regulation in mediating university EFL learners' writing performance. *Innovation in Language Learning and Teaching*, *14*(5), 436–450. https://doi.org/10.1080/17501229.2019.1615493

Teng, L. S., & Zhang, L. J. (2018). Effects of motivational regulation strategies on writing performance: A mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacognition and Learning, 13*(2), 213–240. https://doi.org/10.1007/s11409-017-9171-4

Teng, M. F., Wang, C., & Zhang, L. J. (2022). Assessing self-regulatory writing strategies and their predictive effects on young EFL learners' writing performance. *Assessing Writing*, *51*, 100573. https://doi.org/10.1016/j.asw.2021.100573

Tabachnick, B.G., & Fidell, L.S. (2014). *Using multivariate statistics* (6th ed.). Pearson Inc.

Thomas, N., & Rose, H. L. (2018). Do language learning strategies need to be self-directed? Disentangling strategies from self-regulated learning. *TESOL Quarterly, 53*(1), 20–32. https://doi.org/10.2307/3586262

Tilley, E., & Roach, K. (2017). Ipsative learning: A personal approach to a student's PBL experience within an integrated engineering design cornerstone module. In G. Hughes (Ed.), *Ipsative assessment and personal learning gain: Exploring international case studies* (pp. 129–148). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_7

Towell, R. (2007). Complexity, accuracy and fluency in second language acquisition research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning and teaching* (pp. 461–473). https://doi.org/10.1093/applin/amp048

Tsedev, D., & Purev, E. (2024). Exploring the issues in developing alternative assessments for foreign language courses. *Mongolian Journal of Foreign Languages and Cultures, 28*(596), 1–15. https://doi.org/10.22353/MJFLC2024104

Univio, D. J., & Pérez, A. D. (2019). Ipsative assessment of essay writing to foster reflection and self-awareness of progress. In E. White & T. Delaney (Eds.), *Handbook of research on assessment literacy and teacher-made testing in the language classroom* (pp. 157–180). IGI Global. https://doi.org/10.4018/978-1-5225-6986-2.ch009

Uysal, H. H. (2010). A critical review of the IELTS writing test. *ELT Journal*, *64*(3), 314–320. https://doi.org/10.1093/elt/ccp026

Villarreal, I., & Gil-Sarratea, N. (2020). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research, 24*(6), 874 – 897. https://doi.org/10.1177/1362168819829017

Vorobej, M. (2009). *A theory of argument.* Cambridge University Press. Accessed 18 July 2024, from: https://link.springer.com/article/10.1007/s10503-008-9125-8

Weigle, S.C. (2002). *Assessing writing.* Cambridge University Press. https://doi.org/10.1017/CBO9780511733299

Winstanley, C. (2017). Supporting student learning with cumulative coversheets. In In G. Hughes (Ed.), *Ipsative assessment and personal learning gain: Exploring international case studies* (pp. 65–84). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_4

Wistner, B., Sakai, H., & Abe, M. (2009). An analysis of the Oxford Placement Test and the Michigan English Placement Test as L2 proficiency tests. *Bulletin of the Faculty of Letters, Hosei University, 58*(2), 33–44. http://www.academia.edu/download/

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). L2 development in writing: Measures of fluency, accuracy, & complexity. *Studies in Second Language Acquisition, 23*(3), 423–425. https://doi.org/10.1017/S0272263101263050

Yabukoshi, T. (2021). Self-regulation and self-efficacy for the improvement of listening proficiency outside the classroom. *The Language Learning Journal*, *49*(1), 27–40. https://doi.org/10.1080/09571736.2018.1472626

Yeni, M., Maizarah, M., & Ridhwan, M. (2021). Students' self-regulated learning in English classroom. *ELT-Lectura*, *8*(1), 84–89. https://doi.org/10.31849/elt-lectura.v8i1.5596

Zhang, Z. V., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. *Assessing Writing, 58*, 100784. https://doi.org/10.1016/j.asw.2023.100784

Zhou, J., & Zhang, J. (2017). Using ipsative assessment to enhance first-year undergraduate self-regulation in Chinese college English classrooms. In In G. Hughes (Ed.), *Ipsative assessment and personal learning gain: Exploring international case studies* (pp. 43–64). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56502-0_3