



## **Interactionist and Interventionist Dynamic Assessment Approaches to Teaching Argumentative Writing: Do Complexity, Accuracy, and Fluency Develop?**

**Ayda Rahmani<sup>1</sup>, Mojgan Rashtchi<sup>2\*</sup>, Masood Yazdanimoghadam<sup>3</sup>**

<sup>1</sup> *Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran*

<sup>2</sup> *TEFL Department, Faculty of Foreign Languages, North Tehran Branch, Islamic Azad University, Tehran, Iran*

<sup>3</sup> *TEFL Department, Faculty of Foreign Languages, Garmsar Branch, Islamic Azad University, Garmsar, Iran*

\*Corresponding author: [m\\_rashtchi@iau-tnb.ac.ir](mailto:m_rashtchi@iau-tnb.ac.ir)

(Received: 2021/10/1; Accepted: 2021/2/7)

Online publication: 2021/4/2

### **Abstract**

Complexity, accuracy, and fluency (CAF) are measures of language proficiency in productive skills. This study with a non-equivalent control group pretest-post-test design explored whether Dynamic Assessment (DA) procedures enhances EFL teachers' writing CAF. Three 22-member groups of homogeneous teachers were randomly assigned to one control group (Non-DA), an interactionist DA group (IA-DAG), and an interventionist DA group (IV-DAG). After the nine-session treatment, the study groups took a writing post-test. A one-way ANOVA test showed that the experimental groups outperformed the control group, although no significant differences were found between DA groups. Besides, no statistically significant differences were found between the groups' writing complexity. Considering accuracy, IA-DAG outperformed Non-DAG whereas no significant differences were observed between the experimental groups and between IV-DAG and Non-DAG. Fluency measures showed the superiority of DA groups to the control group. The first conclusion was that DA processes have a decisive role in improving writing skills and the CAF triad. Second, CAF components do not develop in a linear process, and employing DA procedures can be useful for its improvement. The paper has implications for language teachers and SLA researchers.

**Keywords:** feedback, mediation, teaching writing, writing proficiency, EFL teachers

## **Introduction**

Dynamic Assessment (DA) has its origins in Vygotsky's sociocultural theory and invests in the teacher's mediation to improve learners' Zone of Proximal Development (ZPD). The interventionist and interactionist approaches to DA proposed by [Lantolf and Poehner \(2004\)](#) refer to two kinds of intervention (mediation) that intend to solve the problems learners encounter in learning. However, each approach adopts a different way to treat such issues. Interventionist DA focuses on mediation established on pre-scripted hints from implicit to explicit. The interactionist approach invests in learners' responsivity for mediation; thus, it is sensitive to the examinee.

Views on CAF as multi-faceted and multi-layered constituents interrelated in complex ways have changed them into intriguing research subjects ([Larsen-Freeman, 2009](#); [Norris & Ortega, 2009](#); [Skehan, 2009](#)) though have led to indefinite assumptions ([Housen & Kuiken, 2009](#)). Such conclusions are partially due to the conceptualizations regarding their features and how they are related. Complexity, according to [Michel \(2017\)](#), refers to the "size, elaborateness, richness, and diversity of the L2 performance" while accuracy measures "the target-like and error-free use of language" (p.2). However, fluency refers to the "smooth, easy and eloquent production of speech with limited numbers of pauses, hesitations, or reformulations" (p.2). Regarding cognitive processing, an increase in complexity and accuracy indicates the advanced structure of the L2 user's interlanguage. In contrast, a higher fluency designates ease of access to second language knowledge and shows that the L2 user has developed mastery for the automatic use of language ([Skehan, 2009](#)).

One significant finding of the studies on CAF development is that the constituents do not follow a predictable order ([Larsen-Freeman, 2006](#)). Their varying developmental rate among learners evokes the importance of taking note of individual differences envisaged as the vital principle of DA. The sociocultural basis of DA justifies teachers' mediation adjusted to learners' characteristics and needs. Considering the context and society as determinants in learners' language development gives way to employing DA principles in L2 instruction.

Multiple studies have focused on the impacts of DA on the writing skill. For example, [Mauludin and Ardianti \(2017\)](#) and [Alavi and Taghizadeh \(2014\)](#) found that DA positively impacted the participants' writing skills. [Rahimi, Kushki, and Nassaji \(2015\)](#) showed that the interactionist DA could significantly affect L2 writing skills. [Tabatabaee, Alidoust, and Sarkeshikian \(2018\)](#) reported the superiority of cumulative group DA to interventionist DA in improving accuracy in the writings of L2 learners.

Studying the comparative effects of the two procedures in the writing skills has been addressed by a few studies previously; however, not on writing CAF. [Khodabakhsh, Abbasian, and Rashtchi \(2018\)](#) explored the interventionist and interactionist DA models' roles on EFL learners' language awareness level and metacognitive strategy use while engaged in writing activities. The results showed no significant difference between the two DA groups, although they outperformed the control group regarding language awareness. Additionally, [Rahmani, Rashtchi, and Yazdanimoghaddam \(2020\)](#) showed the interventionist DA's superiority to the interactionist and the control groups on writing post-tests of the participants. However, they found no significant differences in the transcendence test administered within a time interval, although both groups outperformed the control group.

According to Vygotsky's Sociocultural Theory (SCT), human activities take place in a cultural context. In his terms, learning is not the one-way flow of knowledge from the more knowledgeable to the less. Learning is a dynamic process in which not only learners but also teachers are involved. Error correction is a social activity involving joint participation and meaningful transactions between the learner and teacher. By mediation, the teacher guides learners to become aware of their errors and correctly reformulate what they want to communicate. Simultaneously, the teacher evaluates learners' responses to the mediation they receive to tailor future instruction to their emerging communicative abilities.

In interactionist approaches to DA, the mediator uses un-scripted mediation forms, calibrated to a learner's emerging needs. Mediation begins with implicit moves to promote learners to find or correct their errors, but it is not pre-scripted. However, interventionist DA consists of scripted mediation in the form of prompts, hints, and questions arranged from the

most implicit to the most explicit ([Lantolf & Poehner, 2004](#)) . Pre-scripted prompts are used as mediation to assess learners' emerging linguistic capacities based on the frequency and quality of prompts needed for correct reformulation. Therefore, mediation in the interventionist approach is not as flexible as the interactionist DA. However, it is less demanding on the mediator, can be used in one-on-one or whole-class interactions, and is more effective in supporting L2 learning than providing feedback non-systematically.

The current study employed the two approaches examining CAF development in writing. Complexity is the most controversial constituent of the triad ([Norris & Ortega, 2009](#)). [Housen, Kuiken, and Vedder \(2012\)](#) consider two dimensions of cognitive and linguistic complexity. They define cognitive complexity as a subjective concept that refers to the "relative difficulty with which language elements are processed during L2 performance and L2 learning" (p. 4). Factors such as aptitude, motivation, and learners' L1 govern this type of complexity. Linguistic complexity, however, embraces "the intrinsic formal or semantic-functional properties of L2 elements (e.g., forms, meanings, and form-meaning mappings)" (p. 4).

The importance of accuracy, the other constituent of CAF, seems to arise from its role in elucidating the path learners follow to learn a second language([Foster & Wigglesworth, 2016](#)). However, its measurement involves some degree of "personal judgment" (p. 112). For measuring this dimension of L2 use, it is necessary to make vital decisions about the norm to choose and the severity of deviance from this norm. However, as [Kuiken and Vedder \(2014\)](#) put forth, rater characteristics may affect their decisions on different writing features, including accuracy.

Fluency is mainly the criterion for measuring proficiency in spoken language. However, studies on writing have also considered fluency a feature that displays writing knowledge ([Wigglesworth & Storch, 2009](#)). [Tavakoli and Skehan \(2005\)](#) discuss three sub-dimensions for fluency; speed (number of words per minute), silence (amount, location, and duration of filled pauses), and repair (false starts, repetitions, and self-corrections). In terms of language processing, speed is associated with control of and access to procedural knowledge; silence reflects the planning and conceptualization

stages of language production, and repair is seen as an indicator of monitoring processes.

All in all, CAF can show the trajectory of L2 development via productive skills. However, the inconclusive findings ([Housen et al., 2012](#)) reported in the literature justify further studies on CAF. Additionally, the effect of personal differences, context, and level of proficiency on the development of CAF stimulates further exploration of their improvement. The existence of a trade-off between the triad pointed out by research findings ([Rashtchi & Yousefi, 2017](#); [Skehan, 2009](#)) encourages further studies on finding out how the components relate. These issues were stimulating enough for the researchers to examine EFL teachers' writing CAF via DA classroom practices. Following [Lantolf \(2005\)](#), the researchers of the present study believed that Dynamic Assessment (DA) has the potentials to enhance writing complexity, accuracy, and fluency (CAF) of EFL teachers within the classroom context. They postulated that writing DA could bridge the gap between individual and group ZPD ([Poehner, 2009](#)) and could develop language proficiency mirrored in the writing skills. This study is significant since, in SLA, CAF is associated with the degree of mastery in L2.

The researchers of the present study presumed that focus on the impacts of interactionist and interventionist DA could add to the understandings of CAF and how these features develop via essay writing. Novice EFL teachers usually encounter the insufficiency of in-service training courses regarding the CAF triad in writing performance. Finding a solution to this issue can lead to a profound understanding of the language constituents and improve their language proficiency level. Therefore, the purpose of the study was to examine whether employing two DA types could increase the CAF of the argumentative writings of EFL teachers. The results might help teachers who look for ways to build the writing proficiency of their students. It also gives some suggestions regarding enhancing grammatical knowledge (accuracy) and lexical knowledge (fluency) via writing.

The objectives of the study led the researchers to formulate the following research questions:

**RQ1:** Is there any difference between the impacts of interactionist DA, interventionist DA, and non-DA approaches on the argumentative writings of EFL teachers?

**RQ2:** What is the impact of interactionist, interventionist, and non-DA approaches on the CAF of EFL teachers' argumentative writings?

## Method

### Participants

The 66 participants (19 males and 47 females) were selected from among 96 Iranian EFL teachers who had graduated in TEFL from Chalus, Tonekabon, and Nour Branches of the Islamic Azad University. The results of a general English proficiency test (GEPT) at the beginning of the course revealed that the scores of 66 individuals were one standard deviation above and below the mean. Thus, they formed a homogeneous group of participants in terms of English language ability with the age range of 25 to 30. IELTS band score calculator indicated that they were at a B2 (intermediate) level of language proficiency equal to 5.5 on the IELTS exam. The band score calculator is accessible at ([https://www.examenglish.com/IELTS/IELTS\\_Band\\_Scores.html](https://www.examenglish.com/IELTS/IELTS_Band_Scores.html)).

The participants were randomly assigned to two experimental groups of interventionist DA group (IV-DAG), the interactionist DA group (IA-DAG), and the control group (Non-DAG). Those individuals whose scores were beyond one standard deviation from the mean took part in other TTC classes appropriate to their proficiency level.

### Instruments

The first instrument used in the study was GEPT, adopted from [CambridgeIELTS11 \(2016\)](#) to control the participants' language proficiency level before the treatment. The test consisted of listening and reading sections (40 questions each) and a writing section. Cronbach's alpha revealed that the listening and reading sections enjoyed a high-reliability index ( $r = .89$ ).

The next tool was a 530-word writing pretest on "*Physical punishment of children must be banned. How far do you agree or disagree?*" (adopted from [Saniei, 2014](#)) that estimated the groups' argumentative writing ability before the treatment. Graduate school admission essays usually include 500-1000 words. Thirty words were considered as fluctuations in counting the number of words. The same topic was used as the post-test to provide the

researchers with a clearer picture of the participants' development on the writing skills in general and CAF in particular. For assessing the participants' writings, the researchers used the Scoring Guide for Writing (2002) developed by California State University Fresno (Appendix A). The inter-rater reliability indices between the two raters who scored the writing pretest ( $r=.98$ ) and post-test ( $r=.99$ ) were computed via Pearson's  $r$ .

The researchers followed [Wigglesworth and Storch \(2009\)](#) for measuring CAF. For complexity, the proportion of clauses to t-units was calculated. For accuracy, global units "expressed in terms of the proportions of error-free t-units of all t-units and error-free of clauses of all clauses" were computed (p. 450). Writing fluency was measured by "calculating the average number of words, t-units, and clauses" (p. 449). Two raters practiced the procedure by rating three essays together to reach an agreement. Inter-rater reliability indices computed via Pearson's  $r$  showed consistency between the two raters (.98 for complexity, .99 for accuracy, and .99 for fluency).

### **Materials**

The researchers prepared nine writing tasks for each treatment session. The content of the tasks was adapted from the IELTS Cambridge writing books. The tasks required the participants to practice writing subordination and coordination, giving color to writing through adjectives, adverbs, using boosters and hedges in writing sentences, writing a five-paragraph essay in about 530 words, based on the given chains of information. The instructor, who was one of the researchers, corrected the tasks but did not score them since their purpose was to help the learners practice their newly acquired knowledge (Appendix B). Also, nine video files performed by native English teachers whose content matched with the course syllabus helped the teacher teach writing. The films are available at [www.engvid.com](http://www.engvid.com).

### **Data Collection Procedure**

The study took 11 sessions. The first session was allocated to the administration of GEPT. During the nine-session treatment (90 minutes each), the three groups performed the writing tasks. The participants' writings were corrected in their presence after the classes to help them observe the process of error correction and ask questions if there were any.

The procedure for error correction usually took 35 to 45 minutes. The participants sat for the 80-minute non-dynamic post-test in the 11<sup>th</sup> session.

### **Control Group (Non-DAG)**

The Non-DAG group received explicit instruction concerning the content of the task to be accomplished. The instructor did not intervene in the process of writing. However, she provided explicit corrective feedback after the participants had completed the tasks. For example, the only piece of information they received on the task for writing a good 'opening sentence' in argumentative essay writing was:

*The task you are supposed to do is writing a suitable opening sentence in English writing. As you see on the slide, there are different tips that you can apply. Based on the information, write down as many opening sentences as you can think of (for the given topic).*

### **Interventionist DA Group (IV-DAG)**

The IV-DAG group received instruction through standardized prompts and feedback, which were arranged from implicit to explicit mediational moves or pre-planned prompts given to participants when they could not perform the intended task. The number and type of regulatory moves in mediation were based on each individual's level of ZPD. They were exposed to the teacher's intervention and assistance in the process of task completion to write the correct answers. The teacher applied concurrent DA ([Poehner, 2009](#)) in which interactions shifted between primary and secondary interactants as one learner's question or comment set the stage for another learner's contribution. Following is an example of the interaction between the mediator and participants (P1, P2, P3).

1. M: Look at your conclusions. See if they look good.
2. P1: No answer.
3. M; For drawing conclusions, what do we need?
- 4: P2: discourse markers.
5. M: Good. We need discourse markers.
4. P1: Like what?
5. M: Like "no wonder," "obviously," or "therefore."  
[The mediator writes a sample sentence on the board].

M: Please check if you have used discourse markers for drawing conclusions.

6. P3: Therefore, [I have]. Therefore, if one wants to be healthy....

7. M: good.

### **Interactionist DA Group (IA-DAG)**

The participants in the IA-DAG group received non-standardized prompts and feedback based on their responsivity. First, the members received instruction regarding the content of the task they had to complete. Then, as they were doing the task, the teacher gave them corrective feedback in the form of non-standardized prompts based on how the participants responded to her feedback. Their responsivity formed the type of feedback they received. Similar to IV-DAG group, concurrent DA was employed in this group. Following is an example indicating the interaction between a participant (P) and the mediator (M).

1. P: We know that 'before to come' is not correct.

2. M: Very good. How should it be corrected?

3. PI: 'before come' or ...

4. M: 'before come' is not a good substitute because 'before' is a preposition, and after it, we cannot use bare infinitive.

5. PI: 'oh, right.' We must put 'ing' after 'before.'

6. M: That's it.

7. PI: Before coming.

8. M: Exactly.

## **Results**

Initially, the researchers examined the participants' homogeneity regarding general English proficiency. The means of IA-DAG ( $M= 22.04$ ,  $SD= 1.55$ ), IV-DAG ( $M= 51.90$ ,  $SD=1.54$ ), and Non-DAG ( $M= 52$ ,  $SD=1.41$ ) groups were very close. The skewness ratios, falling within  $\pm 1.96$  (.346 for IA-DAG, .385 for IV-DAG, and .296 for Non-DAG), showed that the distributions of the scores were normal and running parametric tests was legitimate.

The result of the Levene's test for GEPT showed that the variances enjoyed the assumption of the homogeneity,  $F(2,63) = 0.20$ ,  $p=.81$ , and thus running a one-way Analysis of Variance (ANOVA) was legitimate. The

results of the ANOVA indicated no statistically significant differences among the groups,  $F(2, 63) = .04$ ,  $p = .954$ ) regarding proficiency level (Table 1).

Table 1  
*ANOVA for GEPT*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.212	2	.106	.047	.954
Within Groups	142.773	63	2.266		
Total	142.985	65			

In the next step, the means of the groups obtained from the writing pretest were compared using a one-way ANOVA to ensure that the participants were at the same writing ability level. The Levene's test verified the homogeneity of the variances,  $F(2, 63) = .24$ ,  $p = .85$ . The results of ANOVA (Table 2), showed no statistically significant differences among the groups,  $F(2, 63) = .438$ ,  $p = .647$ .

Table 2  
*ANOVA for Writing Pretest*

		Sum of Squares	Df	Mean Square	F	Sig.
Pretest	Between Groups	45.091	2	22.545	.438	.647
	Within Groups	3245.182	63	51.511		
	Total	3290.273	65			

Table 3 shows the descriptive statistics related to CAF before the treatment. The complexity scores for Non-DAG ( $M = 1.43$ ,  $SD = .211$ ), IV-DAG ( $M = 1.51$ ,  $SD = .275$ ), and IA-DAG ( $M = 1.53$ ,  $SD = .224$ ) were close to each other.

Table 3  
*Descriptive Statistics for CAF, Pretest*

	Groups	N	Mean	SD
Pre-complexity	Non-DAG	22	1.43	.211
	IV-DAG	22	1.51	.275
	IA-DAG	22	1.53	.224
Pre-accuracy	Non-DAG	22	1.04	.274
	IV-DAG	22	1.20	.304
	IA-DAG	22	1.13	.277
Pre-fluency	Non-DAG	22	11.80	1.99
	IV-DAG	22	13.91	2.84
	IA-DAG	22	13.07	1.95

As illustrated in Table 3, the accuracy scores for Non-DAG (1.04, SD=.274), IV-DAG (1.2, SD= .304), and IA-DAG (1.13, SD=.277) also seemed close. The fluency measures for Non-DAG (M= 11.80, SD= 1.99), IV-DAG (M= 13.91, SD=2.84), and IA-DAG (M=13.07, SD=1.95) showed that the control group had a lower mean than the experimental groups. Inferential statistics were used to examine whether the differences were statistically significant.

Table 4 shows three one-way ANOVA test results performed on the CAF of the participants' writing pretests. The results of the pre-complexity,  $F(2, 63) = 1.004$ ,  $p = .372$ ; pre-accuracy,  $F(2, 63) = 2.038$ ,  $p = .139$ ; and pre-fluency,  $F(2, 63) = 2.234$ ,  $p = .116$  showed no statistically significant differences between the CAF of the groups at the outset.

Table 4  
ANOVA for Writing CAF, Pretest

Test		Sum of Squares	df	Mean Square	F	Sig.
Pre-complexity	Between Groups	.115	2	.057	1.004	.372
	Within Groups	3.596	63	.057		
	Total	3.711	65			
Pre-accuracy	Between Groups	.299	2	.149	2.038	.139
	Within Groups	4.614	63	.073		
	Total	4.913	65			
Pre-fluency	Between Groups	21.347	2	10.674	2.234	.116
	Within Groups	301.030	63	4.778		
	Total	322.377	65			

For answering the first research question, descriptive statistics were computed. The skewness ratios of the Non-DAG, IV-DAG, and IA-DAG groups (-.944, -.615, -1.491, respectively), falling between  $\pm 1.96$ , revealed the normality of the distributions. The mean scores of the Non-DAG ( $M=57.61$ ,  $SD=6.77$ ), IV-DAG ( $M=62.88$ ,  $SD=7.15$ ), and IA-DAG ( $M=66.20$ ,  $SD=6.73$ ) groups shows a higher mean score for the IA-DAG group in the post-test.

The one-way ANOVA run on the writing post-test (Table 5) indicates a statistically significant difference among the groups,  $F(2, 63) = 8.69$ ,  $p < .001$ .

Table 5  
ANOVA for Writing Post-test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	825.848	2	412.924	8.695	.000
Within Groups	2991.761	63	47.488		
Total	3817.610	65			

The post hoc Scheffe test traced where precisely the differences laid. Table 6 demonstrates a significant difference between the mean scores of the Non-DAG and IV-DAG and between the Non-DAG and IA-DAG groups. In contrast, no significant difference is observed between the treatment groups.

Table 6  
Scheffe Test, Writing Post-test

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Non-DAG	IV-DAG	-5.27273*	2.07777	.047	-10.4819	-.0635
	IA-DAG	-8.59091*	2.07777	.001	-13.8001	-3.3817
IV-DAG	Non-DAG	5.27273*	2.07777	.047	.0635	10.4819
	IA-DAG	-3.31818	2.07777	.286	-8.5274	1.8910
IA-DAG	Non-DAG	8.59091*	2.07777	.001	3.3817	13.8001
	IV-DAG	3.31818	2.07777	.286	-1.8910	8.5274

To answer the second research question, the researcher(s) ran three one-way ANOVA tests. As Table 7 shows, the mean scores of Non-DAG (M=

1.48, SD= .18), IV-DAG (M= 1.63, SD=.26), and IA-DAG (M=1.62, SD=.19) for complexity do not show any improvement from the pretest to the post-test. However, the mean scores of Non-DAG (M=1.09, SD=.28), IV-DAG (M=1.3, SD= .28), and IA-DAG (M=1.32, SD=.21) for accuracy manifest an increase from the pretest to the post-test. The same results apply to fluency from the pretest to the post-test, Non-DAG : M= 12.87, SD=2; IV-DAG: M= 15.01, SD= 2.88; IA-DAG: M= 14.80, SD=2.32.

Table 7  
*Descriptive Statistics for CAF, Post-test*

	Groups	N	Mean	SD	Std. Error	Min.	Max
Post-complexity	Non-DAG	22	1.48	.186	.039	1.14	1.90
	IV-DAG	22	1.63	.262	.056	1.25	2.10
	IA-DAG	22	1.62	.192	.041	1.33	2.00
Post-accuracy	Non-DAG	22	1.09	.274	.060	.70	1.62
	IV-DAG	22	1.30	.304	.066	.89	1.90
	IA-DAG	22	1.32	.277	.046	.90	1.72
Post-fluency	Non-DAG	22	12.87	2.00	.426	9.40	16.90
	IV-DAG	22	15.01	2.88	.614	10.25	20.00
	IA-DAG	22	14.80	2.32	.495	11.25	19.20

As shown in Table 8, the F value [F (2,63) =2.95, p=.059] does not indicate any statistically significant differences between the groups' post-complexity. However, the measures of post-accuracy [F (2,63) =4.54, p=.014], and post-fluency [F (2,63) =5.12, p=.008] denote significant differences between the groups.

Table 8  
ANOVA Tests for Comparing CAF, Post-test

		Sum Squares	of Df	Mean Square	F	Sig.
Post-complexity	Between Groups	.278	2	.139	2.959	.059
	Within Groups	2.958	63	.047		
	Total	3.236	65			
Post-accuracy	Between Groups	.675	2	.338	4.547	.014
	Within Groups	4.679	63	.074		
	Total	5.354	65			
Post-fluency	Between Groups	61.052	2	30.526	5.166	.008
	Within Groups	372.265	63	5.909		
	Total	433.317	65			

The Scheffe tests were conducted to find precisely where the differences among the groups occur (Table 9).

Table 9  
Scheffe Test for CAF, Post-test

Dependent Variable	Group (I)	Group (J)	Mean Difference (I- J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Post-complexity	Non-DAG	IV-DAG	-.14136	.06534	.105	-.3052	.0224
		IA-DAG	-.13364	.06534	.132	-.2974	.0302
	IV-DAG	Non-DA	.14136	.06534	.105	-.0224	.3052
		IA-DAG	.00773	.06534	.993	-.1561	.1715

	IA-DAG	Non-DA	.13364	.06534	.132	-.0302	.2974
		IV-DAG	-.00773	.06534	.993	-.1715	.1561
Post-accuracy	Non-DAG	IV-DAG	-.20364	.08217	.053	-.4096	.0024
		IA-DAG	-.22409*	.08217	.030	-.4301	-.0181
	IV-DAG	Non-DAG	.20364	.08217	.053	-.0024	.4096
		IA-DAG	-.02045	.08217	.970	-.2265	.1855
	IA-DAG	Non-DAG	.22409*	.08217	.030	.0181	.4301
		IV-DAG	.02045	.08217	.970	-.1855	.2265
Post-fluency	Non-DAG	IV-DAG	-2.14000*	.73293	.018	-3.9775	-.3025
		IA-DAG	-1.92318*	.73293	.038	-3.7607	-.0857
	IV-DAG	Non-DAG	2.14000*	.73293	.018	.3025	3.9775
		IA-DAG	.21682	.73293	.957	1.6207	2.0543
	IA-DAG	Non-DAG	1.92318*	.73293	.038	.0857	3.7607
		IV-DAG	-.21682	.73293	.957	-2.0543	1.6207

As Table 9 shows, there is no statistically significant difference between the Non-DAG, IV-DAG, and IA-DAG regarding complexity in the writing post-test. No statistically significant difference is observed between the writing accuracy of Non-DAG and IV-DAG and between IV-DAG and IA-DAG. However, IA-DAG has outperformed Non-DAG in writing accuracy.

Regarding fluency, there is a statistically significant difference between Non-DAG with the experimental groups. Still, the Scheffe test does not show statistically significant differences between the two experimental groups. That is to say, both types of DA could enhance the post-test writing fluency of the participants.

Partial eta squared was computed to discover what proportion of the variance in the dependent variable is explained by the independent variable. As Table 10 shows, the strength of the relationship between the treatments (interventionist and interactionist) and accuracy and fluency of the writing post-test was strong enough, with the treatment factor accounting for 17.4%, 13.9%, and 14.1% of the variance in accuracy and fluency (Cohen, 1988). Thus, the difference observed between the sample means was large enough to be attributed to the differences between the population means. The reported effect size for the dependent variable post-complexity is medium, with the treatment factor accounting for 9.8% of the variance in the dependent measure.

Table 10  
*Between-Subjects Effects for Effect Size of Dependent Variables*

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Post-complexity	.278 <sup>b</sup>	2	.139	2.959	.059	.098
	Post-accuracy	.675 <sup>c</sup>	2	.338	4.547	.014	.139
	Post-fluency	61.052 <sup>d</sup>	2	30.526	5.166	.008	.141
Intercept	Post-complexity	164.826	1	164.826	3.510E3	.000	.982
	Post-accuracy	101.606	1	101.606	1.368E3	.000	.956
	Post-fluency	13368.516	1	13368.516	2.262E3	.000	.973

Group	Post-complexity	.278	2	.139	2.959	.059	.098
	Post-accuracy	.675	2	.338	4.547	.014	.139
	Post-fluency	61.052	2	30.526	5.166	.008	.141
Error	Post-complexity	2.958	63	.047			
	Post-accuracy	4.679	63	.074			
	Post-fluency	372.265	63	5.909			
Total	Post-complexity	168.062	66				
	Post-accuracy	106.960	66				
	Post-fluency	13801.833	66				
Corrected Total	Post-complexity	3.236	65				
	Post-accuracy	5.354	65				
	Post-fluency	433.317	65				

### Discussion

The affirmative answer to the first research question verified the impact of interactionist and interventionist DA on enhancing the participants' writing ability with no indication of the superiority of one of the approaches to the other. Some studies have reported the benefits of DA-oriented procedures to non-DA procedures in teaching writing ([Alavi & Taghizadeh, 2014](#); [Mauludin & Ardianti, 2017](#)), and some have pointed out the supremacy of

either of the approaches to non-DA methods of teaching writing ([Rahimi et al., 2015](#); [Tabatabee et al., 2018](#)). However, due to different procedures involved in implementing the interactionist and interventionist DA, few studies have focused on their differential impacts on writing. The findings of the present study are notable since they draw on the results from comparing the two DA approaches and a control group in writing classes. The lack of significant differences finds support from [Rahmani et al. \(2020\)](#). They found that both methods showed positive impacts on the learners' writing development in the long term. [Khodabakhsh et al. \(2018\)](#) also found no significant differences between the two approaches though they did not mainly focus on the participants' writing skills. The lack of distinction between interactionist and interventionist approaches in writing is not consistent with the studies conducted on other language skills ([Ahmadi Safa & Beheshti, 2018](#)), which have shown the outperformance of the interactionist DA.

One reason for the efficacy of DA to non-DA approaches relates to the nature of the writing skill. Writing requires the employment of thinking skills such as problem-solving, decision-making, and reasoning ([Rashtchi, 2007, 2019](#)). On the other hand, DA relies on stimulating learners' cognitive abilities to promote their ZPD ([Vygotsky, 1978](#)). Thus, it can be postulated that DA in any of its forms encourages learners' cognitive processing and facilitates focusing on the skills required for writing as well as providing argumentations as tokens of thinking ability. In line with [Fani and Rashtchi \(2015\)](#), it can be asserted that DA procedures direct teachers to realize learners' learning styles and their individualized needs for developing their L2 proficiency.

The second research question addressed the development of CAF in the participants' essays. As the results revealed, there were no statistically significant differences between the groups' writing complexity after the treatment. Regarding accuracy, although IA-DAG outperformed Non-DAG, no significant differences were found between IA-DAG and IV-DAG and between IV-DAG and Non-DAG. However, both DA groups outperformed Non-DAG concerning fluency in writing. A general look into the findings portrays an inconsistency in CAF development, which can imply three conceptualizations. First, it acknowledges the assumptions that CAF is

multi-layered and multi-faceted as the development of the layers needs different degrees of attention and diverging types of instruction. Put differently, each component should be addressed separately. Second, the development of one does not necessarily lead to the development of the other layers. Third, for EFL learners, the constituents of CAF have differing degrees of difficulty. Thus, although complexity, accuracy, and fluency are inter-related, each feature's development is independent of the others. Likewise, the arguments on the limited attentional resources, which posit that focus on one component of CAF leads to neglecting the other constituent or what is known as the trade-off hypothesis ([Skehan, 2009](#); [Vercellotti, 2017](#)), lends support to the findings of the current study.

Moreover, the absence of significant differences regarding the complexity among groups could be attributed to the participants' English language proficiency level. As [Larsen-Freeman \(2006\)](#) and [Wolfe-Quintero et al. \(1998\)](#) reported, the increased written complexity could be witnessed in the written accounts of the more proficient learners' writing skill. Another possible explanation for the insignificant difference in the essays' complexity feature can be related to the type of feedback, which mostly centered on error correction rather than grammatical complexity. That is to say accurate writing received more attention than complex writing in the teacher's mediations.

The statistical results derived from accuracy computations are attention-grabbing. The first inference is that IA-DAG was more effective than Non-DAG but not more effective than IV-DAG. The results become weird when it is noticed that IV-DAG and Non-DAG do not show significant differences. One factor to consider is that the p-value (0.53) indicates that the presence of difference was marginally rejected. Another interpretation gives way to the view that individual variations have a vital role in the changes of the triad ([Housen & Kuiken, 2009](#)).

Additionally, the priority of interactionist DA for increasing accuracy leads the researchers to draw on studies on form-focused instruction and postulate that addressing grammatical errors and drawing learners' attention to form while they are involved in meaning-focused activities are useful in enhancing learners' writing accuracy. Through negotiating errors via

mediation, the teacher provided the learners with the opportunity to "notice and attend to learnable language features in the input" ([Newton, 2013](#)). Thus, the teacher's mediation with each individual and according to the level of his/her level of understanding seems to be more useful than manipulating instruction, which does not consider learners' individualized needs. The quality of the offered mediations in DA groups, which mostly centered on error correction and provided learners with the required feedback, helped them identify and correct their inappropriate written output.

The superiority of the fluency factor in experimental groups to the control group indicates the efficiency of DA. Following ([Skehan, 2009](#)) researchers assume that fluency is a distinct feature in language performance, associated with vocabulary storage. Thus, the point the researchers put forward here is that DA procedures facilitate access to "mental lexicon" ([Skehan, 2009](#)) and help EFL participants retrieve the varied vocabulary they need during writing tasks. Another issue to consider is the preeminence of interactionist DA to interventionist DA as it helps learners overcome the problem of losing accuracy at the expense of fluency proposed by ([Ellis, 1994](#)). In sum, in line with [Housen and Kuiken \(2009\)](#), the findings on CAF imply that complexity is a matter of L2 proficiency and thus more difficult to achieve while accuracy and fluency are related to the learners' interlanguage.

To conclude, the findings support the efficiency of DA processes to non-dynamic methods for developing argumentative writing of EFL teachers and improving CAF. The results obtained highlighted the fact that enhancing the constituents of the CAF triad occurs in a non-linear process, and classroom practices emerging from DA procedures have the potential to stimulate participants' attention to errors. Mediation and interaction also encourage them to look for variety in the use of lexis. The findings underscore the role of interaction and mediation in providing the participants with opportunities to delve into their writing problems and exhibit enhanced performance. The study has implications for EFL instructors whose interest is teaching writing courses. The research shows that writing can contribute to the development of language proficiency when CAF is focused. The findings can be illuminating for researchers interested in the theories related to second language learning.

This study was limited because it did not use qualitative data collection procedures to clarify the participants' underlying reasons for neglecting the complexity factor in their writings. Likewise, the study did not employ techniques to assess how participants could benefit from the mediations and what strategies they used to alter teacher feedback to output.

Further studies on learner strategies that can lead to the emergence of complexity in L2 writing performance are necessary. Additionally, prying into the type of errors that persist in dynamic environments compared to non-dynamic situations can help teachers and educators learn about the potentials of DA in improving writing accuracy.

**Declaration of interest:** none

### References

- Ahmadi Safa, M., & Beheshti, S. (2018). Interactionist and interventionist group dynamic assessment (GDA) and EFL learners' listening comprehension development. *Iranian Journal of Language Teaching Research*, 6(3,Special Issue), 37-56. <https://doi.org/10.30466/IJLTR.2018.120600>
- Alavi, S. M., & Taghizadeh, M. (2014). Dynamic assessment of writing: The impact of implicit/explicit mediations on L2 learners' internalization of writing skills and strategies. *Educational Assessment*, 19(1), 1-16. <https://doi.org/10.22099/JTLS.2017.3956>.
- Cambridge IELTS11. (2016). *Cambridge English IELTS: General training with answers: Authentic examination papers from Cambridge English Language Assessment*. Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press. <https://doi.org/978-0194371896>
- Fani, T., & Rashtchi, M. (2015). Dynamic assessment of reading comprehension ability: Group or individualized. *Education Journal*, 4(6), 325-331. <https://doi.org/10.11648/j.edu.20150406.11>
- Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98-116. <https://doi.org/10.1017/S0267190515000082>

- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, 30(4), 461-473. <https://doi.org/10.1093/applin/amp048>
- Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (Vol. 32). John Benjamins.
- Khodabakhsh, S., Abbasian, G.-R., & Rashtchi, M. (2018). Incorporation of dynamic assessment models into developing language awareness and metacognitive strategy use in writing classes. *Journal of Modern Research in English Language Studies*, 5(4), 55-79. <https://doi.org/10.30479/jmrels.2019.10826.1353>
- Kuiken, F., & Vedder, I. (2014). *Raters' decisions, rating procedures and rating scales*. Sage.
- Lantolf, J. P. (2005). Sociocultural and second language learning research: An exegesis. In *Handbook of research in second language teaching and learning* (pp. 359-378). Routledge. <https://doi.org/10.1558/japl.1.1.49.55872>
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics and Professional Practice*, 1(1), 49-72. <https://doi.org/10.1558/japl.1.1.49.55872>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics*, 27(4), 590-619. <https://doi.org/10.1093/applin/aml029>
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, 30(4), 579-589. <https://doi.org/10.1093/applin/amp043>
- Mauludin, L. A., & Ardianti, T. M. (2017). The role of dynamic assessment in EFL writing class. *Metathesis: Journal of English Language, Literature, and Teaching*, 1(2), 82-93. <https://doi.org/10.31002/metathesis.v1i2.468>
- Michel, M. C. (2017). Complexity, accuracy and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50-68). Routledge. <https://doi.org/9781138936232>.
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17(2), 164-187. <https://doi.org/10.1177/1362168812460814>

- Norris, J. M., & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied linguistics*, 30(4), 555-578. <https://doi.org/10.1093/applin/amp044>
- Poehner, M. (2009). Group dynamic assessment for foreign language pragmatics. *TESOL Quarterly*, 43, 471-492. <https://doi.org/10.1002/j.1545-7249.2009.tb00245.x>
- Rahimi, M., Kushki, A., & Nassaji, H. (2015). Diagnostic and developmental potentials of dynamic assessment for L2 writing. *Language and Sociocultural Theory*, 2(2), 185-208. <https://doi.org/10.1558/lst.v2i2.25956>
- Rahmani, A., Rashtchi, M., & Yazdanimoghaddam, M. (2020). Employing tasks to improve Argumentative Essay Writing of EFL Teachers: A Case of Interactionist versus interventionist dynamic assessment. *Iranian Journal of English for Academic Purposes*, 9(1), 57-75.
- Rashtchi, M. (2007). A pathway toward critical thinking through cooperative writing in an English college course in Iran. *Near and Middle Eastern Journal of Research in Education*, 2 (1), 1-11. <https://doi.org/10.5339/nmejre.2007.2.1>
- Rashtchi, M. (2019). Scaffolding argumentative essay writing via reader-response approach: a case study. *Asian-Pacific Journal of Second and Foreign Language Education*, 4(1), 1-17. <https://doi.org/10.1186/s40862-019-0078-2>
- Rashtchi, M., & Yousefi, L. M. (2017). Reading input flooding versus listening input flooding: Can they boost speaking skill? *Journal of Language and Cultural Education*, 5(1), 39-58. <https://doi.org/10.1515/jolace-2017-0003>
- Saniei, A. (2014). *Raising Iranian EFL learners' pragmatic awareness of intercultural rhetoric in writing through group dynamic assessment* .(Unpublished doctoral dissertation), Islamic Azad University, Research and Science Branch, Tehran, Iran. .
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, 30(4), 510-532. <https://doi.org/10.1093/applin/amp047>
- Tabatabaee, M., Alidoust, M., & Sarkeshikian, A. (2018). The effect of interventionist and cumulative group dynamic assessments on EFL learners' writing accuracy. *Applied Linguistics Research Journal*, 2(1), 1-13. <https://doi.org/10.14744/alrj.2018.36854>
- Tavakoli, P., & Skehan, P. (2005). 9. Strategic planning, task structure and performance testing. In *Planning and task performance in a second*

- language* (pp. 239-273). John Benjamins.  
<https://doi.org/10.1075/llt.11.15tav>
- Vercellotti, M. L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied linguistics*, 38(1), 90-111. <https://doi.org/10.1093/applin/amv002>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes* (E. Rice, Ed. & Trans.). Harvard University Press. <https://doi.org/978-0674576292>
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466. <https://doi.org/10.1177/0265532209104670>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: measures of fluency, accuracy, & complexity*. Second Language Teaching & Curriculum Center, University of Hawaii at Manoa. <https://doi.org/10.1017/S0272263101263050>

### Appendix A

#### Fresno Scoring Guide for Writing

Scoring Level	Knowledge of Conventions	Clarity and Coherence	Rhetorical Choices
4-Accomplished	Besides meeting the requirements for a "3," the writing is essentially error-free in terms of mechanics, models, style, and format appropriate to the assignment.	In addition to meeting the requirements for a "3," writing flows smoothly from one idea to another. The writer has taken pains to assist the reader in following the logic of the ideas expressed.	In addition to meeting the requirements for a "3," the writer's decisions about focus, organization, style/tone, and content made reading a pleasurable experience. Writing could be used as a model of how to fulfill the assignment.

<p>3 -Competent</p>	<p>While there may be minor errors, the paper follows normal conventions of spelling and grammar throughout and has been proofread. Appropriate conventions for style and format are used consistently throughout the writing sample. Demonstrates thoroughness and competence in documenting sources; the reader would have little difficulty referring back to cited sources.</p>	<p>Sentences are structured, and words are chosen to communicate ideas. The sequencing of ideas within paragraphs and transitions between paragraphs makes the writer's points easy to follow.</p>	<p>The writer has made the right decisions about focus, organization, style/tone, and content to communicate clearly and effectively. The purpose and focus of the writing are clear to the reader, and the organization and content achieve the purpose well. Writing follows all the requirements for the assignment.</p>
<p>2-Developing</p>	<p>Frequent errors in spelling, grammar (such as subject/verb agreements and tense), sentence structure, and/or other writing conventions distract the reader. Writing does not consistently follow the</p>	<p>Sentence structure and/or word choice sometimes interfere with clarity. Needs to improve the sequencing of ideas within paragraphs and transitions between paragraphs to</p>	<p>The writer's decisions about focus, organization, style/tone, and/or content sometimes interfere with clear, effective communication. The purpose of writing is not achieved. All</p>

	appropriate style and/or format. Source documentation is incomplete. It may be unclear which references are direct quotes and which are paraphrased.	make the writing easy to follow.	requirements of the assignment may not be fulfilled.
1-Beginning	Writing contains numerous errors in spelling, grammar, and/or sentence structure that interfere with comprehension. Style and/or format are inappropriate for the assignment. Fails to demonstrate thoroughness and competence in the documentation.	Sentence structure, word choice, lack of transitions, and/or sequencing of ideas make reading and understanding difficult.	The writer's decisions about focus, organization, style/tone, and/or content interfere with communication. The purpose of writing is not achieved. Requirements for the assignment have not been fulfilled.

## Appendix B

### Sample Writing Tasks

#### TASK ONE

A. For the topic sentence, write an example and a reason to support it. Look at the sample.

Topic sentence: Physical punishment of children may work in the short-term but not in the long-term.

Example: *For example, my neighbor used to hit their child, and the kid turned into a bully at school and a wife-beater as an adult.*

B. Write a paragraph on the following topic considering the given information.

Topic: Do you think the couple could be successful with opposing religious views?

- Religious differences should be acknowledged.
- Open conversations throughout the relationship about the two religions and what they have in common.
- It depends on how much those religious views have practical implications in their life.

#### TASK TWO

Heart failure and stroke can be caused by a lack of oxytocin released (in the body) as a result of hugging. Following is a list of phrases in a chain that may help you write your essay on *Human beings need to hug and be hugged. Huggers are healthier.* How far do you agree or disagree? You may also use any of the adverbs, connectors, or sequence markers to develop your essay.

*Adverbs:* commonly, likewise, similarly, etc.

*Connectors:* moreover, therefore, consequently, etc.

*Sequence markers:* second, next, finally, etc.

#### Biodata

**Ayda Rahmani** is a TEFL Ph.D. holder from Islamic Azad University, Science and Research Branch in Tehran. She received her B.A. degree in English Literature from Islamic Azad University, Rasht Branch, and M. A. in the same field of study, from Nowrosjee Wadia College, Pune India. In 2015, and got a second master's degree in TEFL from Islamic University, Tonekaboon Branch, after which she was accepted for the Ph.D. program in Tehran. She has been teaching in different language institutes for five years.

**Mojgan Rashtchi** is an associate professor of Applied Linguistics in the faculty of Foreign Languages of Islamic Azad University, North Tehran Branch. She has taught a variety of subjects related to English language teaching to students at different levels. She has published several articles and books and has participated in several local and international conferences. Her primary areas of interest include English language teaching methodology, theories of first and second language acquisition, teaching language skills, and research in education.

**Massood Yazdanimoghaddam** is an assistant professor at IAU, Garmsar Branch. He has been teaching at Allameh Tabatabaei University and different IAU branches for more than two decades at BA, MA, and Ph.D. programs. He has published more than 30 journal articles in both domestic

and international outstanding journals. He is interested in conducting research in Applied Linguistics, especially Pragmatics, Critical Pedagogy, and Translation Studies.